

Portland State University

PDXScholar

Dissertations and Theses

Dissertations and Theses

5-26-1995

Application of a Geographical Information System to Estimate the Magnitude and Frequency of Floods in the Sandy and Clackamas River Basins, Oregon

Dorie Lynn Brownell
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds



Part of the [Geography Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Brownell, Dorie Lynn, "Application of a Geographical Information System to Estimate the Magnitude and Frequency of Floods in the Sandy and Clackamas River Basins, Oregon" (1995). *Dissertations and Theses*. Paper 4877.

<https://doi.org/10.15760/etd.6753>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

THESIS APPROVAL

The abstract and thesis of Dorie Lynn Brownell for the Master of Science in Geography were presented on May 26, 1995, and accepted by the members of the thesis committee and the department.

COMMITTEE APPROVALS:

[REDACTED]

Teresa L. Bulman, Chair

[REDACTED]

D. Richard Lycan

[REDACTED]

Ric Vrana

[REDACTED]

Richard R. Petersen

Representative of the Office of Graduate Studies

DEPARTMENT APPROVAL:

[REDACTED]

Daniel M. Johnson, Chair
Department of Geography

ACCEPTED FOR PORTLAND STATE UNIVERSITY BY THE LIBRARY

by [REDACTED] on *2 Jan 1996*.

ABSTRACT

An abstract of the thesis of Dorie Lynn Brownell for the Master of Science in Geography presented May 26, 1995.

Title: Application of a Geographical Information System to Estimate the Magnitude and Frequency of Floods in the Sandy and Clackamas River Basins, Oregon

A geographical information system (GIS) was used to develop a regression model designed to predict flood magnitudes in the Sandy and Clackamas river basins in Oregon. Manual methods of data assembly, input, storage, manipulation and analysis traditionally used to estimate basin characteristics were replaced with automated techniques using GIS-based computer hardware and software components. Separate GIS data layers representing (1) stream gage locations, (2) drainage basin boundaries, (3) hydrography, (4) water bodies, (5) precipitation, (6) landuse/land cover, (7) elevation and (8) soils were created and stored in a GIS data base. Several GIS computer programs were written to automate the spatial analysis process needed in the estimation of basin characteristic values using the various GIS data layers. Twelve basin characteristic data parameters were computed and used as independent variables in the regression model.

Streamflow data from 19 gaged sites in the Sandy and Clackamas basins were used in a log Pearson Type III analysis to define flood magnitudes at 2-, 5-, 10-, 25-, 50- and 100-year recurrence intervals.

Flood magnitudes were used as dependent variables and regressed against different sets of basin characteristics (independent variables) to determine the most significant independent variables used to explain peak discharge. Drainage area, average annual precipitation and percent area above 5000 feet proved to be the most significant explanatory variables for defining peak discharge characteristics in the Sandy and Clackamas river basins.

The study demonstrated that a GIS can be successfully applied in the development of basin characteristics for a flood frequency analysis and can achieve the same level of accuracy as manual methods. Use of GIS technology reduced the time and cost associated with manual methods and allowed for more in-depth development and calibration of the regression model. With the development of GIS data layers and the use of GIS-based computer programs to automate the calculation of explanatory variables, regression equations can be developed and applied more quickly and easily. GIS proved to be ideally suited for flood frequency modeling applications by providing advanced computerized techniques for spatial analysis and data base management.

APPLICATION OF A GEOGRAPHICAL INFORMATION SYSTEM TO
ESTIMATE THE MAGNITUDE AND FREQUENCY OF FLOODS IN THE
SANDY AND CLACKAMAS RIVER BASINS, OREGON

by

DORIE LYNN BROWNELL

A thesis submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE
in
GEOGRAPHY

Portland State University
1995

ACKNOWLEDGMENTS

I am greatly indebted to those who have come before me, those who surround me and those yet to come... who have, who do, and who shall explore the great complexities of the natural world and share with others their knowledge, their expertise and their enthusiasm. I am fortunate to have benefited from a great wealth of individuals who have shared with me their knowledge, their advice, their encouragement, and expertise, all of which have furthered my education and my appreciation of the physical world that surrounds us, sustains us and gives us life.

My gratitude and respect goes out to the members of my graduate committee; Dr. Teresa Bulman, Dr. Richard Lycan, Ric Vrana and Dr. Richard Petersen. Each has greatly enhanced my knowledge and has been a positive influence during my graduate program at Portland State University. I would also like to thank the Geography department chair, Dr. Daniel Johnson for his advice during the early stages of the project and for his continued enthusiasm and encouragement that carried through to the end.

A special thanks goes to Dr. Teresa Bulman, who kindly took on the challenge of being my primary thesis advisor. I have benefited immensely from her expertise, her advice, and knowledge in the field of water resources, water law and technical report writing. Her teaching skills that demand attention and invite challenge earned my highest respect at the very early stages of my graduate program. I am grateful for having her as an advisor, a professor, and a mentor.

The work herein was prepared in part for the Water Resource Division of the U.S. Geological Survey (USGS) in Portland Oregon under the direction of Dennis Lynch (District Chief). I am greatly indebted to the Portland, Oregon Water Resource Division office for providing the funding and support facilities needed to complete this project. I am especially thankful for the guidance and expertise provided along the way by the many individuals at the USGS. Their input has substantially improved and refined the outcome of the analysis.

Special thanks goes to Antonius Laenen for his technical assistance on the analytical aspects of the regression analysis and for his continued support and guidance from start to finish. I am equally grateful for Howard Harrison's assistance with the Arc Macro Language Programs and the advanced GIS technical support from Mike Darling, Mark Uhrich and Bruce Fisher. Jackie Olson was especially helpful with illustrative and cartographic needs and provided a great source of encouragement and support during many of the difficult phases of the project. Additional thanks goes to Ed Hubbard, Larry Hubbard and John Risley for assistance with peak flow analyses and to Joe Rinella for his supervisory support.

Lastly, there simply are not enough thanks in the world for all those people who are near and dear to me (and you know who you are) —my family (my sister, my mom and my brother-in-law) and my dear friends, who have provided me with an endless supply of patience, encouragement, love and support. Without you, I never would of survived.

TABLE OF CONTENTS

	PAGE
ACKNOWLEDGEMENTS.....	i
LIST OF TABLES	vi
LIST OF FIGURES.....	vii
INTRODUCTION	1
Overview of Geographical Information Systems in Hydrologic Applications.....	5
Overview of Multiple Regression Flood Frequency Analyses	10
Purpose and Scope.....	11
Description of Study Area	12
Physiography and Hydrology.....	12
Land Use.....	14
Climate.....	16
Flood Characteristics	18
Approach.....	19
ESTIMATING THE MAGNITUDE AND FREQUENCY OF FLOODS USING REGRESSION TECHNIQUES.....	21
Data Requirements.....	22
Selection of Gaged Sites	22
Calculation of Flood Magnitudes for Selected Frequencies at Gaged Sites	26
Examination of Factors that Influence Flood Magnitudes.....	27
Selection of Explanatory Climatic and Physical Basin Characteristics	33
APPLICATION OF A GEOGRAPHICAL INFORMATION SYSTEM	37
Computer Hardware and Software Components	37
GIS Data Layer Descriptions	39
USGS Stream Gage Site Coverage	39

Sandy and Clackamas Drainage Basin Coverage.....	40
Stream Gage Subbasin Coverages	40
Hydrography Coverage	40
Precipitation Coverage.....	41
Land Use Coverage.....	42
Elevation Coverage.....	42
Soils Coverage	43
Water Body Coverages	44
Computation of Basin Characteristic Parameters Using GIS Techniques.....	45
Drainage Area.....	47
Average Annual Precipitation	47
Water Storage.....	48
Forest Cover and Glacial Area	48
Main Channel Length.....	49
Main Channel Slope.....	49
Lag Time Ratio.....	50
Total Stream Length	51
Area Above 5000 Feet.....	52
Mean Basin Elevation	52
Soils Index	52
DEVELOPMENT OF MULTIPLE REGRESSION EQUATIONS.....	54
Cross Correlation of Explanatory Variables	55
Models Investigated	59
Two Parameter Model	61
Three Parameter Models	63
Four Parameter Models.....	64
Best Overall Model.....	66
Flood Prediction Accuracy - GIS verses Manual Methods	69
Investigation of Surrogate Variables	71
Discussion	74

SUMMARY OF THE BENEFITS AND LIMITATIONS OF APPLYING A GIS	78
RECOMMENDATIONS FOR FUTURE STUDIES	83
LITERATURE CITED	85
APPENDIX A. PEAK DISCHARGE IN CUBIC FEET PER SECOND FOR SELECTED FLOOD FREQUENCIES AT GAGED LOCATIONS.....	94
APPENDIX B. CONVERSION TABLE.....	95
APPENDIX C. DRAINAGE BASIN CHARACTERISTIC VALUES USED IN THE REGRESSION ANALYSIS AS EXPLANATORY VARIABLES	96
APPENDIX D. GLOSSARY OF SELECTED TERMINOLOGY.....	98
APPENDIX E. ARC MACRO LANGUAGE PROGRAMS.....	103

LIST OF TABLES

TABLE		PAGE
I.	U.S. Geological Survey gaging stations selected for the flood frequency analysis	24
II.	Frequency of use of the various basin characteristics in rural regression equations developed for individual States and Puerto Rico	35
III.	Anderson and others (1976) Level I land use classification scheme	43
IV.	Pearson Product - Moment correlation coefficient matrix of explanatory basin characteristic variables for the Sandy and Clackamas river basins	57
V.	Basin characteristic parameters and range of values for the Sandy and Clackamas river basins, Oregon.....	59
VI.	Comparison of 2-year recurrence interval regression equations using different sub-sets of explanatory variables	64
VII.	Three parameter regression equations for the Sandy and Clackamas river basins	67
VIII.	Comparison of predicted 2-year flood magnitudes computed using regression equations developed using GIS and manual methods	70
IX.	Comparison of three parameter models for the 2-year recurrence interval	74

LIST OF FIGURES

FIGURE	PAGE
1. Sandy and Clackamas river basins, Oregon	13
2. Land use in the Sandy and Clackamas river basins, Oregon	15
3. Spatial distribution of average annual precipitation in the Sandy and Clackamas river basins, Oregon	17
4. Map showing location for 19 U.S. Geological Survey streamflow gaging stations used for the flood frequency analysis and contributing subbasin drainage areas	25
5. Flow diagram showing the GIS data layers and programs used to compute basin characteristic parameters used in the regression analysis.....	46
6. Plot of the residuals of predicted flood magnitudes for the 2-year recurrence interval using the 2, 3, and 4-parameter regression equations.....	62
7. Plot showing the linear relation between the log of Drainage Area and the log of Total Stream Length.....	73

INTRODUCTION

Flood prediction has long been of goal of hydrologists. Federal, state and local agencies responsible for flood management programs depend on accurate and up-to-date flood information to minimize structural damage and loss of life. Because it is not economically feasible to install stream-gaging stations at all places where flood information is needed, alternative methods have been developed to estimate flood magnitudes and frequencies at ungaged stream locations. These methods range from simple flood formulas developed in the early part of the century (Benson, 1962a) to the refinement of complex statistical methods and flow models being used today.

One of the most effective and well established techniques for transferring flood characteristics from gaged to ungaged stream locations is a multiple regression approach. Regression equations are developed using multiple regression techniques that define the relation between flood-peak discharges at gaged locations and the climatic and physical characteristics of the contributing watershed. Once flood frequency relations have been determined at several gaged locations, the established regression equations can then be used to determine flood-peak discharges at ungaged stream locations by using measured values of climatic and physical basin characteristic information. For example, climatic and physical basin characteristics such as contributing drainage area, area of lakes and ponds, percent forest cover, and precipitation intensity are used to predict flood magnitudes at ungaged stream locations in the High Cascade Region in

western Oregon (Harris and others, 1979). Regionalized regression equations for determining the magnitude and frequency of floods for rural, unregulated watersheds have been published for every state in the nation (Jennings and others, 1994). In Oregon, multiple regression techniques have been applied in flood frequency analyses by Lystrom (1970), Harris and others (1979), and Harris and Hubbard (1983).

Regression techniques provide reliable estimates of peak flow statistics and have become the standard approach for computing flood magnitudes at ungaged stream locations; however, the process of assembling and computing the necessary basin characteristic data for the regression analysis requires considerable effort. Flood frequency analyses often rely on manual methods, such as the use of planimeters and transparent grid overlays, to extract basin characteristic and climatic data from a variety of mapped sources that vary in scale and geographic projection. The effort involved in organizing, retrieving, storing, analyzing and manipulating the data has been labor intensive, time consuming and costly since much of the time is spent assembling large volumes of data from hard copy maps, data reports, tabular output, and field notes. Time is also expended in the computational process of aggregating the data (using manual averaging and weighting techniques) into a format that can be used in the regression model.

The process of developing the data necessary for a flood frequency analysis have imposed several limitations in achieving the most accurate results of flood prediction. First, the time, labor, and expense of using manual methods to assemble, analyze, and compute data parameters limits the ability of the analyst to explore all of the data available and to change model parameters as needed. The time expended in the data assembly

process limits the analyst's ability to focus on the development and calibration of the model and presumably, hinders the achievement of the most accurate results.

Second, manual methods are subject to human error. Most historical data bases are non-digital, and the analysts is often required to make qualitative estimates or generalized assumptions about the data using "judgment calls" or "eye-balling" techniques (Berry and Sailor, 1987). This subjective approach to estimating certain model input parameters increases the chance of data inaccuracies and inconsistencies.

Third, manual methods are time-consuming, so updates are not regularly performed. Regional flood frequency analysis require periodic updates to assess the effect of alterations in the hydrologic flow regime due to artificial or natural changes in the landscape and to take advantage of additional climatic and hydrologic data. Updates, however, are often avoided because of the expense involved in re-evaluating new data parameters and re-running the regression analysis. As a result, the lack of periodic updates adversely affects the accuracy of a given study over time.

The goal of providing the most accurate and reliable flood information requires the use of the best tools and information available. The technology capable of automating the process of data extraction and performing rapid analysis of complex spatial data is currently available. Geographical information systems (GIS) are specifically designed to capture, store, manipulate, analyze and display geo-referenced digital data. These systems use hardware and software computer components and are designed to provide digital mapping capabilities and data base management.

In the past fifteen years, the use of GIS has received widespread attention and has been applied in numerous hydrologic applications. Only recently however, has the use of GIS been applied to flood-related studies (Hill and others, 1987, Battaglin and others, 1993, Hay and others, 1993, Woodbury and Jawed, 1993) or multiple regression applications (See and others, 1992). There is a need to implement and take advantage of the advanced capabilities of these spatial analysis tools and apply these to flood frequency analyses. Further advancements in flood prediction demand the exploration of new approaches that can greatly improve the process of examining and understanding complex spatial interactions of hydrologic, climatic and physical processes that define flood characteristics.

Identifying the environmental elements and processes that lead to a flood event involves a holistic approach—an approach that is inherent to physical geography. This approach draws on the information and principles derived from several of the earth sciences such as hydrology, climatology, biology, soil science, geology, ecology and geomorphology. A geographic approach brings together scientific elements of the various earth sciences, examines the spatial inter-relationships and weaves this information into recognizable patterns that are more easily understood. Significant advances in flood hazard research have resulted from the work of geographers who have applied an interdisciplinary approach to flood analysis (Greis and Wood, 1981, Gupta and Fox, 1974, Hirschboeck, 1988, Kates, 1962, Platt, 1986, Waylen, 1985, Waylen and Woo, 1982, White, 1945, 1964, Whyte, 1986, Wolman and Miller, 1960, Woo and Waylen, 1984, 1986).

This thesis uses a geographic approach to performing a flood frequency analysis by synthesizing information from several of the earth science

disciplines, performing spatial analyses among the various environmental elements and identifying the physical and climatic factors within the region that interact and cause flood events. GIS is an essential component to this geographic approach by providing the technology and tools necessary to accomplish the goal of spatial analysis, data synthesis and data output. In the interest of achieving the highest level of accuracy for flood prediction and striving to use the best tools available, it is critical that the next step in the evolution of spatial data analysis be taken by applying GIS technology to flood modeling studies and identifying the benefits and limitations of its application.

OVERVIEW OF GEOGRAPHICAL INFORMATION SYSTEMS IN HYDROLOGIC APPLICATIONS

Since the 1980s, use of GIS technology has evolved from obscurity to common practice in businesses, universities and government agencies throughout the United States and internationally. Within the field of hydrology, GIS has gained widespread use and recognition in several sub-disciplines that encompass a wide range of water-related applications. GIS has helped water resource managers with management decisions regarding water use and consumption (Schoolmaster and Marr, 1992), river basin management (Goulter and Forrest, 1987), stormwater/wastewater management (Cowden, 1991, Hobert, 1989, Meyer and others, 1993), and water resource planning (Berich, 1985, Leipnik and others, 1993, Leipnik and Loaiciga, 1991, Weghorst and others, 1991, Wright and Buehler, 1989, Zelt, 1991).

A survey of GIS applications in surface water hydrology was performed by an American Society of Civil Engineers (ASCE) (1993) task

committee during the fall of 1992. Based on the information supplied by survey questionnaires received from approximately 100 respondents, the ASCE reported that GIS was most frequently being used in areas of planning, modeling, land use, environmental concerns, and natural resource applications. A similar survey was performed by DeVantier and Feldman (1993) who cited the use of GIS in floodplain management, flood forecasting, erosion prediction/control, water quality prediction/control and drainage utility implementations.

A review of GIS applications in flood-related topics has revealed several trends. The most notable is the rapid increase in the number and different types of applications since 1990. Most applications have been tailored to address a specific hydrologic problem within a particular region of interest. Battaglin and others (1993) and Hay and others (1993) performed a hydro-climate study using GIS in the Gunnison River Basin in Colorado. Jeton and Smith (1993) used GIS techniques to develop a watershed model for two Sierra Nevada drainage basins in California. Similar studies have been performed in the Big Sandy River basin in Kentucky (Bhasker and others, 1992), the Bull Run watershed in Pennsylvania (Shamsi, 1993), and the Amite River basin in Louisiana (Hill and others, 1987).

GIS is most widely used to develop digital data layers, perform spatial analyses and compute input data parameters for various types of hydrologic models. Bhasker and others (1992) used GIS to develop data layers representing stream networks, basin boundaries, landuse, soils and strip-mined areas. These data layers were spatially overlaid and used to compute input parameters such as main channel length, drainage density, sinuosity ratio, basin shape and basin area which were input into a

geomorphological instantaneous unit hydrograph model used to compute watershed runoff volumes.

GIS has been used to develop the input parameters needed for the Penn State Runoff Model (PSRM) which is a watershed rainfall-runoff simulation model used to simulate runoff hydrographs for various durations and frequencies (Shamsi, 1993). GIS has also been used for modeling purposes in urban areas for storm-runoff prediction (Berry and Sailor, 1987), and urban storm-water management (Meyer and others, 1993). Stuebe and Johnston (1990) used GIS techniques to develop input parameters needed for a stormwater runoff model using GRASS (Geographic Resources Analysis Support System) software to estimate runoff discharge volumes. In flood modeling applications, GIS is most frequently used to quantify basin characteristics for modeling purposes.

Several flood studies have used the Soil Conservation Service (SCS, recently renamed the Resource Conservation Service) method to estimate storm runoff for a specific drainage basin. The SCS method uses soil type and land use data to compute a SCS *curve number* which is used to estimate the volume of runoff from an entire basin based on a variety of precipitation inputs. Numerous studies have used GIS to develop soil and land use data layers and perform spatial analysis to compute SCS *curve numbers* (Hill and others, 1987, Stuebe and Johnston, 1990, Schmidt and Romak, 1991, Bhaskar and others, 1992, Shamsi, 1993, and Woodbury and Jawed, 1993). For example, Berry and Sailor (1987) used the SCS method to estimate storm runoff for the Race Brook watershed in Connecticut. GIS techniques were used to automate the procedure for calculating input parameters needed for the SCS model by performing spatial overlays of soil and landuse data

themes to compute a SCS *curve number*. SCS *curve numbers* were determined for subbasin units using area-weighted averaging techniques. Muzik and Pomeroy (1990) used the SCS runoff *curve number* and a regional dimensionless unit hydrograph method for predicting flood frequency curves for selected watersheds in Alberta's Rocky Mountain foothills in Canada.

GIS has often been used to model the topography of a region using digital terrain models such as digital elevation models (DEMs) and triangular irregular networks (TINs). These data structures provide a digital representation of the earth's surface using elevation data and are useful for applications such as modeling flow paths of surface water runoff and delineating drainage basin boundaries. Tachikawa and others (1993) used a TIN-DEM to establish direction of water flow, automatically delineate stream networks and determine topographic attributes such as slope, aspect, flow path lengths and contributing drainage areas. Daly and Neilson (1992) and Daly and others, (1994) used DEMs to model the distribution of precipitation in mountainous regions. Digital elevation data are useful for rain-runoff modeling (Beven and Moore, 1993) and obtaining stream parameters such as stream length, stream slope and stream order (Connors and others, 1989, Lorenz, 1990). Digital terrain modeling is useful for obtaining geomorphic, biological and hydrologic parameters of a particular drainage basin (Moore and others, 1993, Silfer and others, 1987, O'Callaghan and Mark, 1984). Computer algorithms have been developed to automatically delineate watershed boundaries using DEMs and TINs (Jenson and Dominique, 1988, Jenson, 1991, Jones and others, 1990, Marks and others, 1984).

Applications of GIS differ in the type of digital data structures used, such as vector or raster-based. The first application of a GIS in hydrologic modeling used a raster (or grid cell) data structure for storing terrain information (Pentland and Cuthbert, 1971). A raster data structure is a matrix made up of equally-sized units called "pixels" or "cells" in which a numeric value is stored. Grid data developed from DEMs contain elevation values within each cell. Complex algorithms and models have been developed using grid data structures to analyze water movement and identify flow paths using terrain modeling techniques. Much of the work involving GIS in hydrologic applications has utilized raster data or cell-based data structures (Beven and Moore, 1993, Connors and others, 1989, Darling and Hubbard, 1994, Lorenz, 1990, Moore and others, 1993, O'Callaghan and Mark, 1984, Silfer and others, 1987, Tachikawa and others, 1993).

Unlike raster formats, vector-based digital formats use points, lines and polygons to represent mapped features. Attribute information, such as stream name and stream length, can be linked to these mapped features through a specifically designed GIS relational data base. This study uses primarily a vector-based digital data format to represent mapped features such as streams, gaging stations and drainage basins. Both vector and raster-based formats of digital data storage have inherent limitations and benefits. DeVantier and Feldman (1993) provide a more detailed discussion on the different types of digital formats used in hydrologic applications.

GIS continues to receive widespread attention and use in the field of hydrology. The appeal of automating the process of geographical analysis and the unique ability to manipulate and display complex data in a simplified format has contributed to the popularity of this technology. The temporal and spatial nature

of water, both above and below the surface, and the fundamental relationship of map information and spatial data to hydrologic analyses, have made hydrology a natural field for the application of GIS.

OVERVIEW OF MULTIPLE REGRESSION FLOOD FREQUENCY ANALYSES

Techniques used in flood prediction are not fixed but rather continue to evolve. Models and methods have varied based on the type of flood information required and have ranged from simple flood formulas to complex statistical methods and flow models being used today. A comprehensive history of the various methods used to determine flood characteristics during the early part of the century (prior to 1936) is provided by Jarvis and others (1936). Benson (1962a) provides a historical overview of the evolution of methods and models used to evaluate the occurrence and magnitude of floods up to the early 1960's.

During the 1960's, Benson (1959, 1962a, 1962b, 1964) made extensive use of multiple regression techniques in several flood investigations. Results of these earlier flood studies and later work by Thomas and Benson (1970), demonstrated that through the use of multiple regression techniques, reliable estimates of flood-peak discharges could be determined for ungaged stream locations based on climatic and physical basin characteristics. By the 1970's, use of multiple regression techniques for relating flood peak-discharges to basin characteristics became the standard approach of the U.S. Geological Survey (USGS) for estimating the magnitude and frequency of floods for rural, unregulated watersheds (Jennings and others, 1994). Since 1973, regression equations for estimating the magnitude and frequency of floods have been published for every state in the nation.

In Oregon, multiple regression techniques have been used in several regionalized flood frequency analyses (Hulsing and Kallio, 1964, Lystrom, 1970, Harris and others, 1979, Harris and Hubbard, 1983). The most recent state-wide flood frequency analysis for the State of Oregon was performed by Harris and others in 1979 for Western Oregon and in 1983 by Harris and Hubbard for Eastern Oregon. Information from these flood frequency analyses provides useful material for performing similar multiple regression techniques in this flood frequency analysis of the Sandy and Clackamas river basins in Oregon.

PURPOSE AND SCOPE

The purpose of this study was to analyze the effectiveness of replacing traditional manual methods of data assembly, input, storage, manipulation and analysis with automated techniques using GIS-based computer software and hardware components. The primary study objectives are to: (1) use a GIS to develop and compute basin characteristics for use in a regression analysis for the prediction of flood magnitudes at ungaged stream locations, (2) compare the level of accuracy for flood prediction using manual and GIS methods, and (3) identify the benefits and limitations of applying a GIS in a flood frequency analysis. The study area was limited to the Sandy and Clackamas river basins in order to provide a manageable study region as a preliminary step to establishing the feasibility of applying a GIS in regionalized flood frequency analyses for larger scale applications, such as statewide flood frequency analyses.

DESCRIPTION OF STUDY AREA

The Sandy and Clackamas river basins encompass an area of approximately 1,440 square miles in northern Oregon (fig.1). The Sandy river basin has an area of about 500 square miles, and the Clackamas basin, located to the south, has an area of 940 square miles. The adjacent drainage basins are bounded on the north by the Columbia River, on the east by the Cascade Range, and on the west and south by the Willamette River Basin.

Physiography and Hydrology

The Cascade Range spans the eastern boundary of the study area in a north-south direction. Mt. Hood (elevation 11,245 feet) is the most notable topographic feature of the region. Stream tributaries in the Sandy River Basin originate high on the west and south slopes of this glacier-capped volcanic peak and flow west through several deep, glacier-carved valleys before eventually joining the Columbia River to the north. The Sandy River has two principle tributaries: the Bull Run River (which is impounded twice by dams forming Bull Run Reservoirs I and II) and the Salmon River. Smaller tributaries include Lost Creek, ZigZag River, Little Sandy River, and Gordon Creek; these tributaries have steep gradients, narrow canyons and numerous waterfalls.

The headwaters of the Clackamas River Basin drain the heavily forested slopes of the Cascade Range between Mt. Hood to the north and Mt. Wilson to the south. Tributaries of the Clackamas River originate at approximately 6000 foot elevation and descend to the valley where the Clackamas River joins the Willamette River. Tributaries of the Clackamas River include the

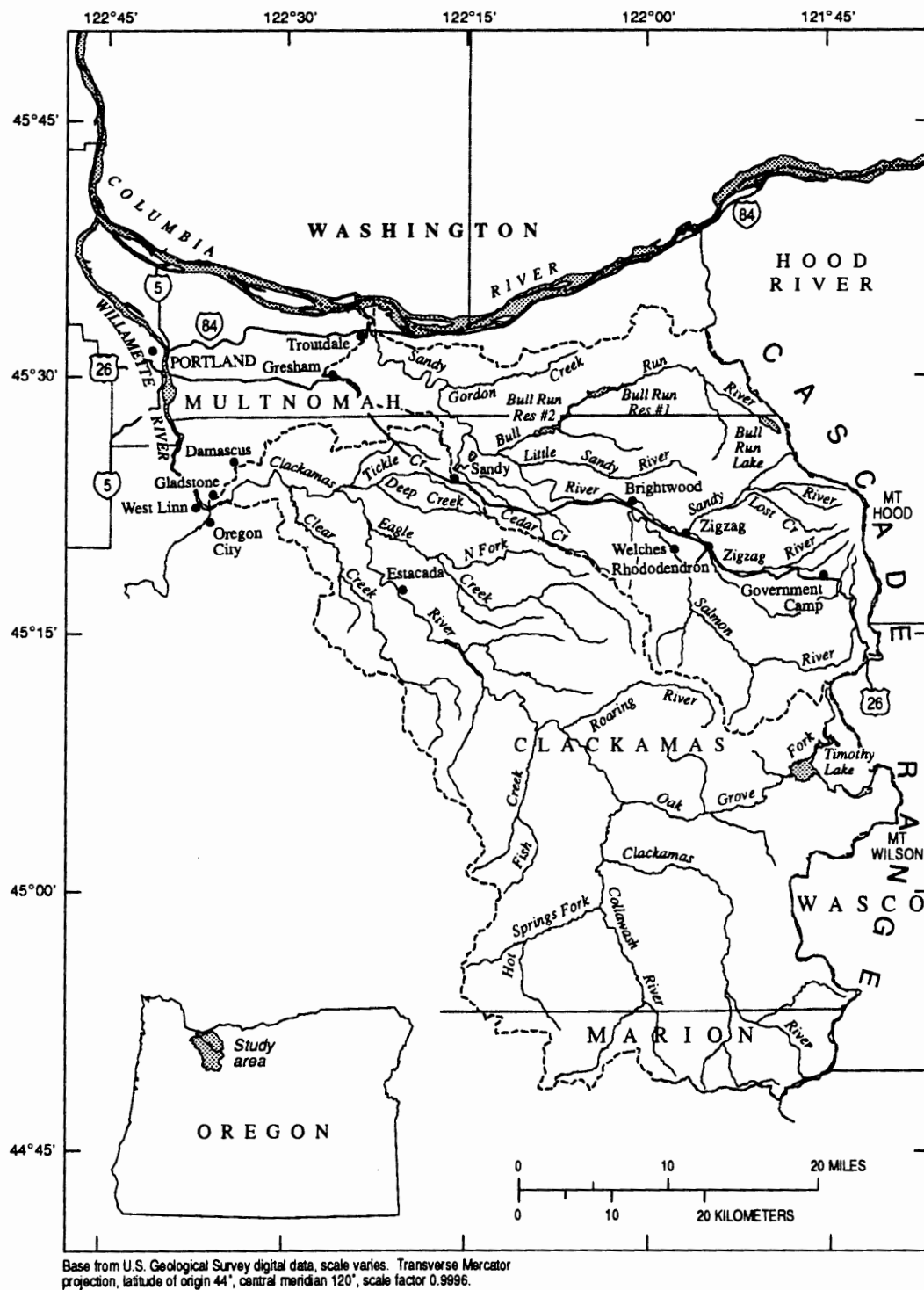


Figure 1. Sandy and Clackamas river basins, Oregon.

Collawash River, Hot Springs Fork, Oak Grove Fork, Roaring River, Fish Creek, Eagle Creek, Deep Creek, Clear Creek and Tickle Creek.

Land Use

Land use (and land cover) information is needed to provide an important overall assessment of a basin's physical characteristics, which in turn provide an indication of the factors influencing runoff processes. The majority of the Sandy and Clackamas basins are forested (fig. 2). Forest cover affects runoff processes by the interception of rain and snowfall and by modifying the accumulation and melting of snow. The remainder of the land cover is divided into agricultural lands, urban areas, tundra, snow fields and glaciers. A very small portion consists of rangeland or wetlands. Perennial snow and glacial fields, located on Mt. Hood, provide temporary storage of surface runoff during the winter months (as precipitation accumulates in the form of snow) and augment surface runoff in the spring as snow- and glacial-melt occurs.

Wetland areas occur throughout the Sandy and Clackamas basins. Several are located in isolated areas, distant from stream channels, in the higher elevations near Mt. Wilson. The effect of wetlands on peak discharge depends largely on their proximity to the stream channels. Isolated wetlands, distant from the stream channel, have minimal or no effect on peak discharge. However, wetlands adjacent to stream channels may significantly reduce flood peaks by providing dispersion and retention areas for flood waters that overflow stream banks.

Impervious surface areas, typical of the urban environment, significantly increase surface water runoff (Laenen, 1978, 1983) and are an

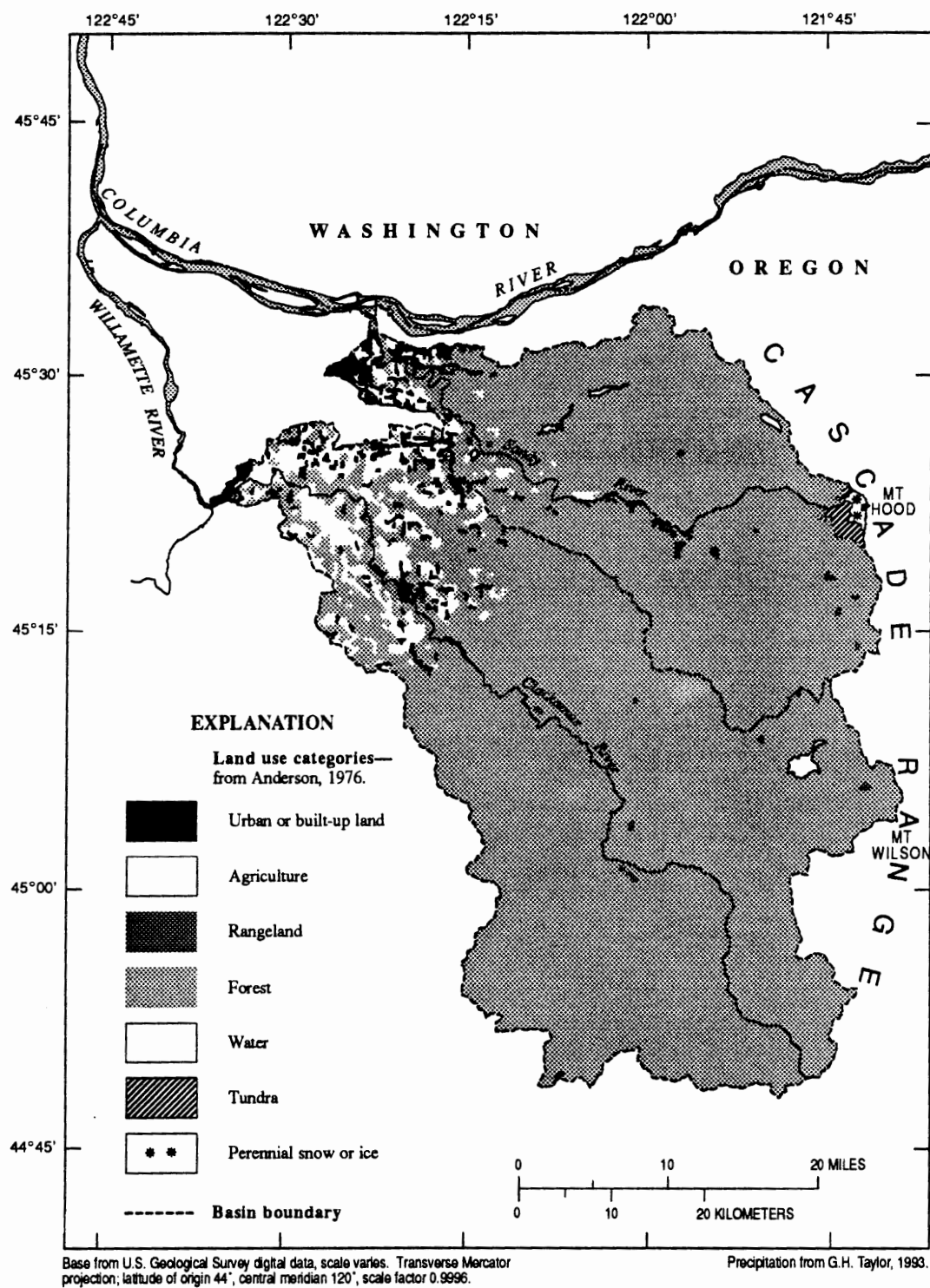


Figure 2. Land use in the Sandy and Clackamas river basins, Oregon.

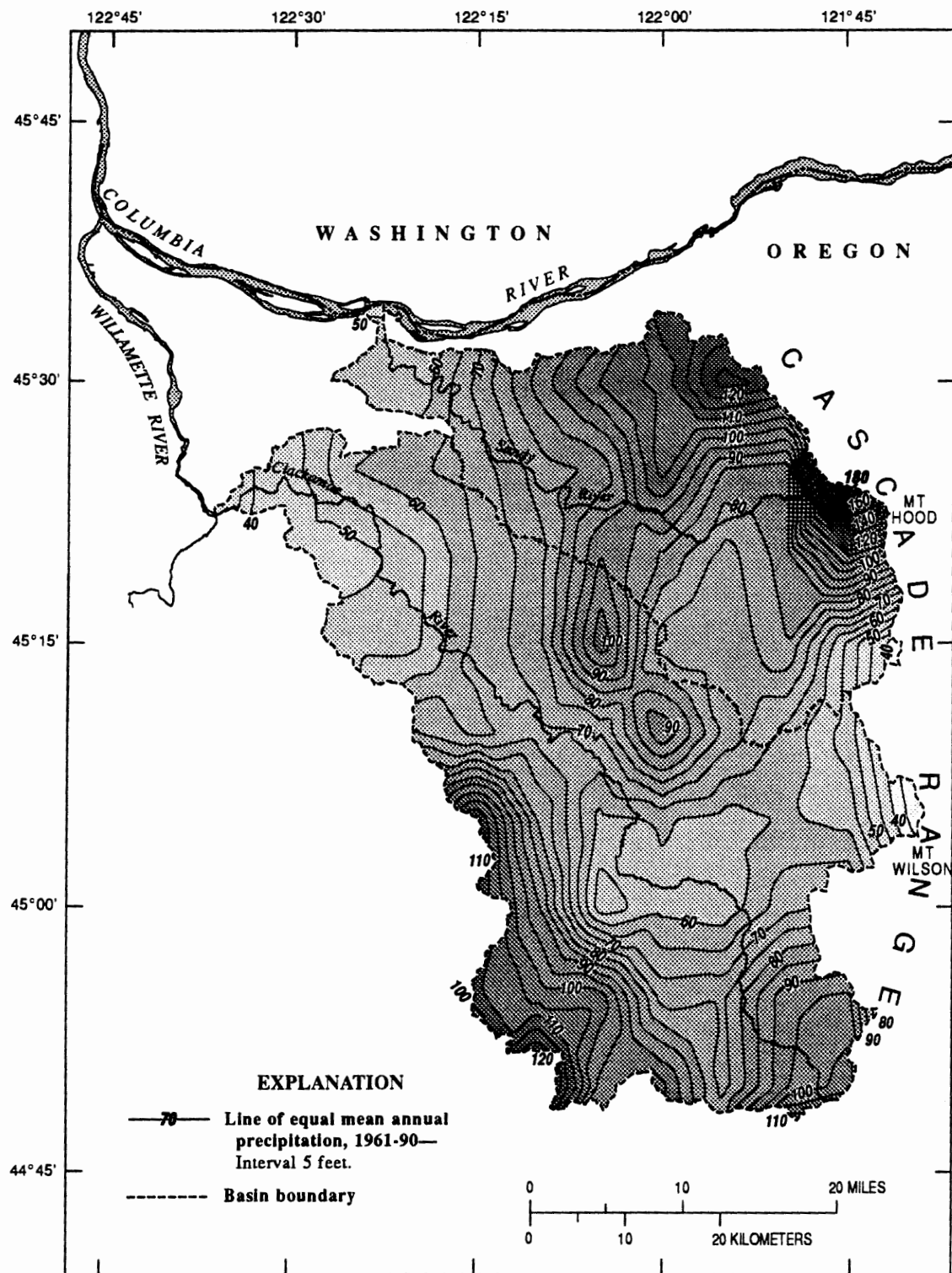
important factor in determining storm runoff volumes for urban areas. The small portion of urban area in the Sandy and Clackamas basins is located on the western most fringe of the study area at the basin outlets. These areas are the outer extensions of the Portland metropolitan area. Because urban areas represent a minimal portion (less than 3 percent) of the overall study area, impervious surface cover was not considered in this analysis.

Climate

Ocean influences tend to moderate the climate in this region and as a result, temperature ranges are not great. The climate is characteristically humid and temperate. Winter months are typically cloudy and wet, whereas summer months are relatively clear and dry. Unusually high temperatures are often the result of warm winds originating east of the Cascade Range.

The Cascade Range greatly influences the climatic and hydrologic regime of the study area by creating a natural topographic barrier to the prevailing westerly winds and marine air-masses that originate from the Pacific Ocean approximately 90 miles to the west. As the moisture-laden air-masses encounter the Cascade Range and are forced upward, cooling and condensation occur causing prolonged periods of precipitation.

Average annual precipitation ranges from 40 inches at the confluence of the Sandy and Columbia rivers to more than 180 inches at the higher elevations around Mt. Hood (Taylor, 1993). The isohyetal map (fig. 3) shows the spatial distribution of precipitation throughout the study region. Precipitation increases with elevation on the windward (or western side) of the Cascade Range, and much of the winter precipitation accumulates as snow at elevations above 3000 feet.



Base from U.S. Geological Survey digital data, scale varies. Transverse Mercator projection: latitude of origin 44°, central meridian 120°, scale factor 0.9996.

Precipitation from G.H. Taylor, 1993.

Figure 3. Spatial distribution of average annual precipitation in the Sandy and Clackamas river basins, Oregon.

Approximately 78 percent of total annual precipitation occurs during the winter months from October to March (Lystrom, 1970) when westerly winds bring in rain producing cyclonic storms. Precipitation during the winter months is characterized as frequent, of long-duration, and caused by low-to-moderate intensity frontal storms.

Flood Characteristics

Flood events west of the Cascade Range commonly occur during two high-water periods each year. From October through March heavy winter rains cause maximum rates of runoff, and from April through June spring flooding often results in the higher elevations when heavy rainfall runoff occurs in combination with runoff from melting snow (Harris and others, 1979). The largest and most destructive floods on record have occurred during the months of December, January and February (Brands, 1947) and were the result of winter storms originating from the southwest. **These winter storms caused unusually high rates of runoff as warm rain reached the higher elevations and fell on previously accumulated snowpack causing excessive snowmelt in combination with heavy rainfall.**

The greatest flood ever recorded for regions near the Sandy and Clackamas river basins occurred in December 1861 when rainfall for the months of November and December was 225 percent and 140 percent (respectively) of normal (Brands, 1947). Prior to the flood-producing storm, above normal-precipitation combined with below-normal temperatures causing excessive accumulations of snowpack in the higher elevations. In the last days of November and the first few days of December, warm south winds and prolonged periods of excessive rainfall created the dangerous

combination of heavy rainfall further augmented by excessive runoff from melting snow.

The "Christmas Flood" of December 1964 remains the largest flood event to have occurred in the area within the past one hundred years. It was the largest in areal extent and was responsible for severe flooding, extensive damage and loss of lives in five nearby states. In Oregon, flooding was caused by intense rains falling on frozen ground and accumulated snowpack resulting in accelerated rates of heavy runoff.

A review of historic flood events occurring within the Western Cascade Region of Oregon provides useful information on the seasonality and the contributing physical and climatic factors that often accompany flood events. Historic flood records can be useful for describing flood characteristics at ungaged locations through the use of regression analysis.

APPROACH

The selection of a study area was based primarily on the availability of GIS digital data for basin characteristics (such as hydrography, land use, soils, water bodies, topography, and precipitation) and the existence of an adequate data base for streamflow statistics. The use of pre-existing GIS digital data would significantly reduce the time spent developing a GIS data base by minimizing the time spent in data assembly and extraction from hard copy maps and data reports. It was equally important to have an adequate stream gaging network that had ten or more years of continuous streamflow records and represented natural streamflow conditions with no (or minimal) upstream diversions or regulation. Based on these

considerations, the Sandy and Clackamas river basins provided a suitable study region for a flood frequency analysis using a GIS.

GIS-based computer software (ARC/INFO version 6.1.1) was used to develop the necessary basin characteristic data parameters and GIS-based computer hardware components replaced manual tools such as planimeters and transparent grid overlays. In order to demonstrate the accuracy of flood predictions using GIS techniques, it was important to choose a region where previous flood studies had been performed with manual methods of data analysis. A set of regionalized regression equations had been previously developed for the Sandy and Clackamas region by Harris and others (1979) in a flood frequency analysis of western Oregon. By comparing the results achieved by Harris and others (1979) using manual methods to those achieved using GIS techniques, it would be possible to determine if the level of confidence for flood prediction was retained using GIS techniques.

Several GIS-based computer programs were developed to automate the process of performing spatial overlay operations among various GIS data layers, area and line-weighted computations, and advanced routing processes which would allow for the computation of basin characteristic values for the regression analysis. These techniques were outlined and evaluated to identify the benefits of using GIS techniques over manual methods and to identify and note any serious pitfalls or limitations. These observations were essential for determining the feasibility of applying GIS techniques in large-scale flood frequency analyses.

ESTIMATING THE MAGNITUDE AND FREQUENCY OF FLOODS USING REGRESSION TECHNIQUES

Multiple regression techniques are used to develop prediction equations for estimating individual values of a dependant variable, such as peak stream discharge, based on several independent variables, such as basin characteristics representing land use, drainage area, elevation, and average annual precipitation. The independent (or explanatory) variables are factors or characteristics considered to be physically related to the dependant (or response) variable. In this case, climatic and physical basin characteristics known to influence flood characteristics were selected and used in the regression analysis as explanatory variables. Flood magnitudes for select recurrence intervals of 2-, 5-, 10-, 25-, 50- and 100-years were computed using streamflow records collected at gaged stream locations and were used as the response variable.

As outlined in Harris and others, (1979), equations used to define the relation between a response variable and several explanatory variables may be expressed by the following mathematical equation

$$Q_T = K C_1^a C_2^b C_3^c \dots C_n^z$$

in which Q_T is the discharge for a selected recurrence intervals (2-, 5-, 10-, 25-, 50- and 100-year), T ; K is a regression constant; C_1 , C_2 , C_3 , and C_n are basin characteristics; and a , b , c , and z are regression coefficients. Various basin characteristic are regressed on select flood magnitudes to determine which basin characteristics best define the response variable (peak

discharge) at gaged stream locations. Once the statistical relation is established at gaged locations, flood magnitudes can be determined for ungaged stream locations simply by computing and inputting the appropriate basin characteristic values into the established prediction equations.

DATA REQUIREMENTS

Data requirements for the multiple regression flood frequency analysis involved the computation of flood magnitudes for selected recurrence intervals at gaged locations and the computation of selected basin characteristic parameters. Several phases of activities were required to develop the necessary data needed for the flood frequency regression analysis of the Sandy and Clackamas river basins. These included: (1) the selection of gaged stream locations having no (or minimal) upstream regulation or diversions and a minimum of ten years of continuous streamflow records, (2) calculation of flood magnitudes for recurrence intervals of 2, 5, 10, 25, 50, and 100 years at selected gaged locations using existing streamflow records and a log-Pearson Type III flood frequency distribution, (3) identification of the physical and climatic factors that influence the magnitude and frequency of floods in the basins, and (4) selection and computation of the necessary drainage basin characteristics.

Selection of Gaged Sites

A suitable streamflow gaging network requires each gaged site to have a minimum of 10 years of annual peak discharge data in order to provide statistical significance (Benson, 1962b). Longer records of peak flow data

provide a stronger basis for statistical inferences for predictions of future flood magnitudes —not in terms of specific events, but in terms of probability of recurrence for selected time intervals.

A suitable gaged site should also have a minimum amount of upstream regulation from reservoir storage or stream diversion. The amount of upstream regulation affects flood frequency analyses when applying regression, correlation or interpolation methods for transferring streamflow data from gaged sites to ungaged locations, because these methods are only applicable where natural streamflow conditions exist.

A network of 19 USGS stream-gaging stations, each having 10 or more years of annual peak flow records and representing subbasins with essentially natural streamflow conditions, was selected for the analysis. The station number, station name, drainage area, period of record and number of years of record for each gaging station are listed in Table I.

Because regression methods are limited to the analysis of streamflow records for gaged sites with no (or minimal) upstream regulation, the period of record for Station 14209000 had to be modified to take into account the construction of Timothy Lake reservoir in 1956. The period of record for Station 14209000 extends from 1909 to the current year, however, only those records collected prior to the construction of Timothy Lake, during the period of 1910 to 1953, were used in the regression analysis.

Annual peak discharge data were collected from 13 gaging stations in the Sandy basin and 6 gaging stations in the Clackamas basin. Contributing drainage areas (subbasins) above each gaging site range in size from 3.8 to 934 square miles, and the number of years of recorded streamflow information ranges from 10 to 82 years. The location of the gaged sites and corresponding station numbers are shown in Figure 4.

TABLE I.

**U.S. GEOLOGICAL SURVEY GAGING STATIONS SELECTED FOR FLOOD
FREQUENCY ANALYSIS.**

STATION NUMBER ¹	STATION NAME	DRAINAGE AREA	PERIOD OF RECORD	NUMBER OF YEARS OF RECORD
SANDY BASIN				
14131000	Little Zigzag at Twin Bridges	3.78	1926-1936	10
14131400	Zigzag River near Rhododendron	14.68	1981-1994	12
14134000	Salmon River near Government Camp	8.00	1926-1987	61
14134500	Salmon River below Linney Creek	52.53	1927-1950	23
14135000	Salmon River at Welches	98.48	1925-1936	11
14135500	Salmon River above Boulder Creek	105.93	1936-1952	16
14137000	Sandy River near Marmot	259.26	1911-1994	82
14138800	Blazed Alder Cr near Rhododendron	8.17	1963-1994	30
14138850	Bull Run River near Multnomah Falls	48.33	1966-1994	27
14138870	Fir Creek near Brightwood	5.34	1975-1994	18
14139700	Cedar Creek near Brightwood	7.90	1965-1994	28
14139800	South Fork Bull Run near Bull Run	15.53	1974-1994	19
14141500	Little Sandy River near Bull Run	23.84	1919-1994	74
CLACKAMAS BASIN				
14208000	Clackamas at Big Bottom	136.10	1920-1970	50
14208500	Oak Grove at Timothy Meadows	52.92	1913-1929	16
14209000	Oak Grove above Powerplant Intake ²	125.88	1910-1953	43
14209500	Clackamas above Three Lynx	488.70	1922-1987	66
14210000	Clackamas River at Estacada	679.87	1908-1987	79
14211000	Clackamas River near Clackamas	933.94	1962-1983	21

¹ U.S. Geological Survey station number based on downstream order.

² Period of record used in the analysis prior to the construction of Timothy Lake.

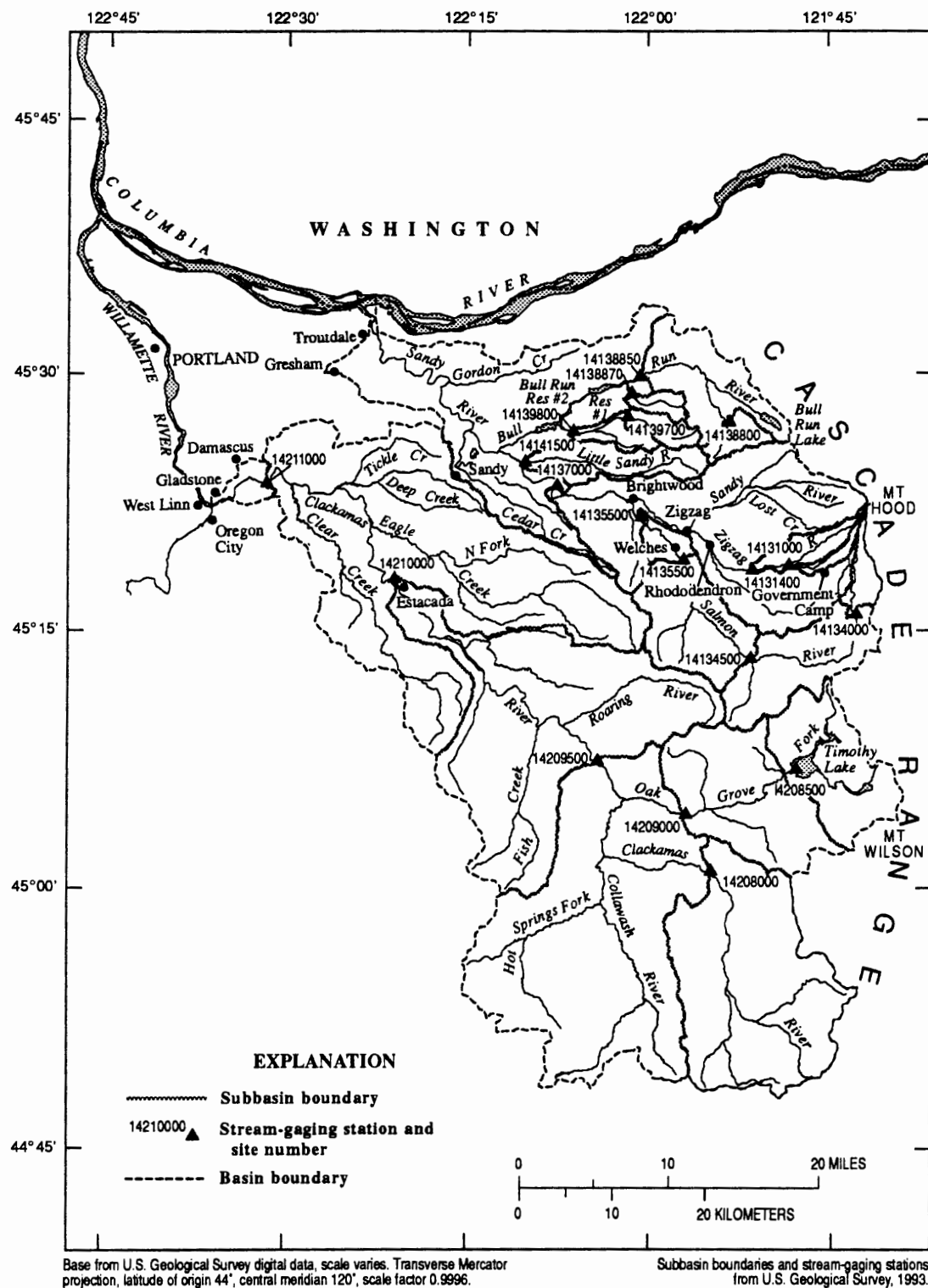


Figure 4. Map showing location for 19 U.S. Geological Survey streamflow gaging stations used for the flood frequency analysis and contributing subbasin drainage areas.

A more comprehensive description of each gaging station can be found in the USGS Statistical Summaries of Streamflow Data Report for Oregon (Wellman and others, 1993).

Calculation of Flood Magnitudes for Selected Frequencies at Gaged Sites

Annual peak flow frequencies for recurrence intervals of 2-, 5-, 10-, 25-, 50- and 100- years were computed for the 19 gaging stations located within the Sandy and Clackamas river basins using guidelines outlined by the Interagency Advisory Committee on Water Data (1982). These guidelines, commonly referred to as Bulletin 17B, outline the procedures for computing annual flood-frequency curves for gaged sites based on the statistical analysis of systematic streamflow records. Flood potential computed for gaged locations is expressed in terms of peak discharge and exceedance probabilities and is determined for gaging stations with a minimum historical record of 10 years of systematic stream-gaging records.

Program J407 was developed by the USGS and is outlined in the USGS WATSTORE manual (Kirby, 1981). The program follows the guidelines of Bulletin 17B to compute a log-Pearson Type III frequency curve based on the mean standard deviation and skewness of the logarithms of the recorded annual peak flows. The log-Pearson Type III distribution, which uses a logarithmic transformation of the flood data, is recommended as the basic distribution for defining the annual flood series (Interagency Advisory Committee on Water Data, 1982). Program J407 outputs a table listing the estimated log-Pearson Type III flood frequency curve for the supplementary systematic record and flood magnitudes for selected exceedance probabilities. The systematic record is weighted by a regional skew based on

record length. This information was retrieved from the USGS Water Data Storage and Retrieval System (WATSTORE) for the 19 USGS gaging stations within the Sandy and Clackamas basins. Flood magnitudes for recurrence intervals of 2-, 5-, 10-, 25-, 50- and 100-years are recorded in cubic feet per second and are presented in Appendix A. A conversion table for converting English units to metric units is provided in Appendix B.

Examination of Factors that Influence Flood Magnitudes

The most common cause of flooding in the Pacific Northwest is short-duration high intensity(convection) rainfall, long-duration low intensity rainfall, and snowmelt events. On the western slopes of the Cascade Range, precipitation and snow melt are the primary causes of flood events. For small basins, landslides or debris flows can also cause peak flows to increase in magnitude.

Understanding flood characteristics requires an understanding of the factors that underlay these events. For example, when precipitation reaches the ground (whether in the form of rain or snow) its distribution, rate of runoff, and direction of flow or storage will be controlled by several factors. Temperature and elevation will affect the type of precipitation, such as rainfall or snowfall. Meteorological factors, such as wind direction, dew point and radiation, influence runoff by affecting evaporation and snow melt rates (Benson, 1962b). Topographic features such as drainage basin area, slope, and type of ground cover will also affect rates and the distribution of runoff. Below is a brief description of several factors that are known to influence runoff characteristics. More detail can be found in Benson (1962b).

Drainage Area. Drainage area is the total contributing basin area located upstream from a particular stream site (such as a gaging station). Earlier regression analyses (Harris and others, 1979, Patterson, 1971, Benson, 1962b, and Thomas and others, 1994) have shown drainage area to be the most statistically significant explanatory variable affecting peak discharge. As expected, runoff rates increase in a downstream direction as small stream channels converge to form larger stream channels. Similarly, the volume of flood waters increase in a downstream direction as water accumulates from larger drainage areas.

Precipitation. ~~Rainfall is the most common direct precursor of flood events.~~ Precipitation varies from region to region, in areal extent, level of intensity, and degree of magnitude. One way of expressing these variations in a manner that bears meaning to flood peaks is by using an average value. Average annual precipitation, for example, provides an good indication of the general climatic regime by providing a index of relative wetness or dryness throughout a region and of the relative magnitude of storm events within the region for specified frequencies.

Specific rainfall intensity can be a more meaningful characteristic in predicting flood events. However, the intensity duration that affects a peak-flow event is dependent on basin size and the individuality of a storm event. If intensity is to be used as an explanatory variable over annual precipitation, then that intensity duration associated with the lag-time of the basin should be used. Use of a singular intensity duration is the same as using an annual intensity because isopluvial distributions exhibit nearly the same patterns (Laenen, 1995).

Slope. Stream channel slope and slope of the surrounding land surface greatly influence the rate of runoff and flow velocity in the stream channel. Steeper slopes, typical of stream tributaries located high up in mountainous areas, cause high streamflow velocities and increase the rate at which water travels across a land surface and within a stream channel. Waterfalls and steep gradients typical of the upper elevations of the Sandy and Clackamas basins provide a clear example of this phenomenon.

Several indices for indicating slope have been tested. However, there is no standard method or universally accepted technique for evaluating channel slope (Langbein, 1947, Benson, 1962b, Harris and Hubbard, 1983). One commonly used index is the "85-10" slope factor (Benson, 1962b), which is used to determine average slope, in feet per mile, of the main channel between points 10- and 85 percent of the distance upstream from the gaging site to the basin divide. Elevation is determined at these points from contour lines on topographic maps.

Water Storage. Previous flood frequency analyses (Harris and others, 1979, Benson, 1962b, Laenen, 1980) have found basin storage to be one of the more significant explanatory variables used for defining flood magnitudes. Flood peaks can be significantly reduced by the retention and temporary storage of surface water runoff by lakes, ponds, wetlands, reservoirs and flood plain areas. Water retention also occurs in surface depressions, stream channels, groundwater and in soil, but the lack of reliable information and difficulty of developing simple numerical indices makes it difficult to evaluate the effect of these storage facilities on the analyses. Laenen (1983) was able to use the surface area of lakes, ponds, marshes, flood plains, depressions and detention-storage facilities to compute storage input values

for regression equations designed to estimate storm runoff magnitudes in urban areas. The inclusion of water storage significantly lowered the average standard error of estimate in that study and improved the level of confidence for runoff prediction within the urban study region.

Elevation. Although elevation does not directly influence flood discharge, it can be used to evaluate other factors that vary with elevation such as precipitation, vegetation, snow melt, evaporation rates, radiation, and temperature. In regions of relatively homogeneous climatic conditions, mean basin elevation provides a good indicator of type of precipitation falling on a basin, such as rain or snow or a combination of the two. This parameter is easily obtained from digital elevation models currently available through the U.S. Geologic Survey National Mapping Division at various map scales.

Elevation provides an important indicator of the presence of ephemeral snow fields that provide temporary storage of precipitation. Elevation is also a useful criterion for determining the transition zone where snow melt occurs by providing a relative index of air temperature. According to estimates provided by the State of Oregon Water Resources Board (1965), approximately 1/3 of the total precipitation at the 4,000 foot elevation, falls as snow and more than 3/4 of the total precipitation at the 7,000 foot elevation falls as snow. For this study, the 5,000 foot elevation was chosen as the transition zone above which precipitation was considered to be in the form of snow. The 5,000 foot index was used previously in a flood analysis by Thomas and Benson (1970).

Snow Melt and Glacial Melt. On the western slopes of the Cascade Range, snow melt combined with excessive rainfall has caused several of the region's largest floods on record. Rainfall augmented by snow melt or glacier melt substantially increases the amount of water for runoff, which results in increased flood peaks. Glacial melt from Mt. Hood is a significant source of streamflow that contributes to seasonal snow melt and supplies runoff when the snow cover has been depleted. It is difficult to obtain accurate information on snow pack or snow and glacier melt processes. Harris and others (1979) indicated that better estimates of peak flow and lower standard errors for their flood-frequency equations could be achieved for the high Cascade Region if this information was available.

Fountain and Tangborn (1985) discuss several modeling techniques for predicting runoff from glacial areas. These modeling techniques are beyond the scope of this analysis but provide important information on glacier and snow melt processes.

Stream Density. Stream density is a measure of the total length of all contributing stream tributaries above a selected stream site per square mile of the total basin area and has been known to influence the timing of flood peaks. Benson (1962b) indicated that stream density did not show any significant relation to peak discharge provided that channel slope and water storage were taken into consideration. Stream density can also be represented by total upstream channel length which is simply the total length of contributing stream channels.

Soils and Geology. ~~Soils and geologic features of a basin influence the size of flood magnitudes by controlling the flow of water as it travels by~~

surface and sub-surface routes. Factors such as soil porosity, infiltration rates, surface permeability, available storage capacity, and transmission rates influence rates of runoff and groundwater flow. Storm runoff in some instances may disappear into fissures and underground channels, in which case geologic features become important considerations since they will reduce the size of floods. These factors are highly related because the geology is the parent material for the soils. Geologic features as well as soil characteristics are often too difficult to evaluate or quantify numerically and are in many cases not considered in flood frequency analyses. Soils may be classed into groups based on infiltration rates, soil porosity, depth, permeability and transmissibility. For example, the SCS groups soils using a hydrologic soil-group classification system which classifies soils as A, B, C or D based on the intake of water "at the end of long-duration storms occurring after prior wetting and opportunity of swelling, and without the protective effects of vegetation" (Benson, 1964).

Forest and Vegetative Cover. The amount and type of vegetative cover influences streamflow and surface runoff by transpiration, interception of precipitation, interception of surface runoff into the vegetative substrate, and by modification of the accumulation and melting of snow. Highly vegetated or heavily forested areas will intercept precipitation and retain large volumes of water and will therefore decrease runoff rates. By comparison, areas that have steep slopes and little or no vegetation will have higher rates of runoff.

The effect of clear-cutting on storm-runoff magnitudes remains a controversial subject. Several studies indicate that the removal of forest

cover increases peak flows and attribute the increase to changes in soil moisture content. Recent studies, however, have not detected a statistically significant increase in the size of peak discharge in areas where clear-cut logging had occurred and soils had not been significantly disturbed (Rothacher, 1971, 1973; Harr and others, 1975; Harr, 1976).

Dunne and Leopold (1978) indicated that infiltration rates are commonly high in humid regions, because the vegetative cover protects the soil from rain-packing and dispersion processes. The presence of humus and organic material provides an open soil structure that allows for the storage of water. It is often difficult to adequately measure the effect of vegetative cover on runoff characteristics.

Additional Factors. The physical, climatic and hydrologic factors that affect flood frequencies and magnitudes are numerous and complex. For example, factors such as stream order, meander and bifurcation ratios, valley width, mean monthly precipitation, average annual evaporation, orientation, thunderstorm days, different indexes of temperature, and basin width, length, and shape have all been indicated as providing useful measures for defining flood characteristics (Thomas and Benson, 1970, Benson, 1962b, 1964). Whereas most provide useful information for defining flood characteristics, many are difficult, if not impossible, to evaluate and express in mathematical terms. In many instances, the data are simply not available or are too costly and time consuming to develop or obtain.

Selection of Explanatory Climatic and Physical Basin Characteristics

Once the physical and climatic factors known to influence flood magnitudes were identified, a subset of these were selected for the multiple

regression analysis. The selection criterion was based on the results of previous flood studies for this region (Lystrom, 1970, Harris and others, 1979) which found certain physical and climatic factors to be more significant than others in explaining flood characteristics. The most recent flood frequency analysis (that included the Sandy and Clackamas basins) by Harris and others (1979) reported drainage area, precipitation intensity, area of lakes and ponds, and forest cover as being the most significant basin characteristic variables for defining flood magnitudes in this region. With the exception of precipitation intensity, values for each of these basin parameters were computed and used in the flood frequency analysis for the Sandy and Clackamas river basins.

Additional climatic and physical factors previously not considered in flood studies for this region by Lystrom (1970) and Harris and others (1979), but believed to be significant and worthy of consideration were also selected for the regression analysis. These were: (a) percent area above 5000 feet, (b) percent glacial cover and (c) total upstream channel length. The availability of newly developed digital data and the use of GIS technology allowed for examination of these additional data parameters previously not considered in flood analyses. It was believed that the inclusion of these additional basin characteristics could improve the regression relation and better explain the variations in flood magnitudes occurring in the Sandy and Clackamas river basins.

The four most frequently used basin characteristics used in regional regression equations developed for unregulated watersheds in the United States and the Commonwealth of Puerto Rico are drainage area, main channel slope, mean annual precipitation, and area of lakes and ponds.

Table II shows the frequency of use of various basin characteristics used in rural regression equations developed throughout the United States and Puerto Rico. Several of these basin characteristics were considered in this flood frequency analysis.

TABLE II.

FREQUENCY OF USE OF THE VARIOUS BASIN CHARACTERISTICS IN
RURAL REGRESSION EQUATIONS DEVELOPED FOR INDIVIDUAL STATES
AND PUERTO RICO.

Basin Characteristics used in Rural Regression Equations	Number of States (including Puerto Rico)
Drainage Area (square miles)	51
Main Channel Slope (feet per mile)	27
Mean Annual Precipitation (inches)	19
Storage/Area of Lakes and Ponds (percent)	16
Rainfall amount for a given duration (inches)	14
Elevation of Drainage Basin (feet)	13
Forest Cover (percent)	8
Channel Length (miles)	6
Minimum Mean January Temperature (degrees F)	4
Basin Shape [$((\text{length})^2 \text{ per drainage area})$]	4
Soil Characteristics	3
Mean Basin Slope (feet per foot or feet per mile)	2
Mean Annual Snowfall	2
Area of Stratified Drift (percent)	1
High Elevation Index (percent basin above 6000 feet)	1
Relative Relief (feet per mile)	1
Drainage frequency (number of first order streams per square mile)	1

Table II was taken from the "Nationwide summary of U.S. Geological Survey regional regression equations for estimating the magnitude and frequency of floods for ungaged sites, 1993" (Jennings and others, 1994).

The following basin characteristics values were computed for each gaging station and were used in the flood frequency regression analysis: 1) drainage area, 2) main channel slope, 3) main channel length, 4) mean elevation, 5) water storage, 6) glacial areas, 7) lag-time ratio, 8) total upstream river length, 9) area above 5000 feet, 10) soil permeability, and 11) forest cover. Computed values for drainage basin characteristic used in the regression analysis are listed in Appendix C.

APPLICATION OF A GEOGRAPHICAL INFORMATION SYSTEM

The following discussion outlines the use of a geographical information system for calculating climatic and physical characteristics of a watershed to be used in regression modeling to estimate flood magnitudes. A description of the computer software and hardware components as well as the development, source and techniques used to create the GIS data base and climatic and physical basin characteristic variables for the Sandy and Clackamas flood frequency analysis are provided below. A glossary of selected GIS and statistical terminology is provided in Appendix D.

COMPUTER SOFTWARE AND HARDWARE COMPONENTS

ARC/INFO, a GIS-based computer software system available through Environmental Systems Research Institute, Inc. (ESRI, 1990, 1991, 1992) was used to assemble and develop the GIS digital data layers and perform automated spatial analysis procedures necessary to compute the explanatory basin characteristic variables. ARC and INFO are the two subsystems that make up the GIS computer software system. ARC contains the utilities used to create, manipulate, analyze, and graphically display spatial data. INFO is a relational data-base management system (DBMS) utilized through ARC which stores attribute information such as map boundaries, tic registrations, feature attribute information, and source documentation that is associated with the individual data layers. INFO files contain feature attributes, such as an Arc Attribute Table (AAT) or a Polygon Attribute Table (PAT) which allow

attribute information, such as stream name or stream length, to be stored in a tabular file directly linked to a corresponding data layer.

Version 6.1.1 of the ARC/INFO software system (current at the time of the project) was loaded on a Data General AViiON 300 workstation platform using a UNIX operating system. The workstation was equipped with sixteen megabytes of random access memory and all the data were stored and processed within a 400,000 kilobyte project directory mounted on a local disk.

Several ARC Macro Language (AML) programs were written to automate the process of computing explanatory basin characteristic input variables. The AML programs organize and group ARC/INFO commands into a series of instructions which the computer uses to perform multiple geo-processing operations such as computing the length (in miles) of the longest stream reach located upstream of a selected gaged stream location. The GIS computer programs used in this report are presented in Appendix E and are described in the section entitled "Computation of Explanatory Basin Characteristic Variables".

Digital data layers were acquired in an ARC/INFO format from local, state and federal agencies. Data layers are often digitized at varying map scales and map projections and as a result, each data layer received was checked for accuracy and projected into a single transverse mercator geographic map projection. The transverse mercator projection is a conformal projection that has a constant scale along any chosen central meridian and differs from the regular mercator projection which has a constant scale along the Equator (Snyder, 1982). Since central meridians run in a north-south direction, the transverse mercator projection is most useful

for regions having a dominant north-south extent. The transverse mercator projection was selected to minimize the distortion of all properties within 15 to 20 degrees on either side of the $-120^{\circ} 00' 00''$ central meridian which included the Sandy and Clackamas river basins.

GIS DATA LAYER DESCRIPTIONS

The digital data layers used to compute the necessary basin characteristic input parameters for the regression analysis were 1) the location of the 19 USGS gaging stations, 2) the subbasin boundaries delineating contributing drainage area above the 19 gaged sites, 3) hydrography of the Sandy and Clackamas basins, 4) isohyetal contours representing average annual precipitation, 5) land use, 6) elevation, 7) soil types, 8) water bodies and wetlands at a 1:100,000 map scale and 9) water bodies and wetlands at a 1:24,000 map scale. The use of the term *data layer* and *coverage* are used interchangeably.

USGS Stream Gage Site Coverage

The USGS Stream Gage Site Coverage is a single data layer showing the spatial distribution and location of the 19 USGS gaging stations used in the regression analysis. The digital data layer was created by retrieving latitude and longitude values from the USGS Water Resource Data Report (Hubbard and others, 1994) and recording these values into an American Standard Code for Information Interchange (ASCII) flat file format. Coordinates (recorded in degrees, minutes, seconds) were projected into a Universal Transverse Mercator (UTM) coordinate system and then generated as a points coverage using ARC/INFO's 'GENERATE' command that creates a

digital map containing point features representing the location of each of the 19 gaging stations. The data coverage layer was then projected from UTM to transverse mercator and attribute items, such as station number and station name, were added to the related attribute file (Point Attribute Table) in INFO.

Sandy and Clackamas Drainage Basin Coverage

This data coverage contains the basin boundaries of the Sandy and Clackamas river drainage basins. It served as an essential component for building the GIS digital data base by providing a template to obtain and create additional digital data layers. The outer basin boundary was used as a "cookie-cutter" to extract digital data for the Sandy and Clackamas basins from a number of statewide digital data coverages containing digital information on hydrography, elevation, land use, water bodies, precipitation and soils.

Stream Gage Subbasin Coverages

Contributing drainage areas above the 19 gaged stream sites were manually drawn on USGS 1:24,000 scale topographic maps, and were digitized and stored as polygon features in a single data layer. From this data layer, 19 separate subbasin boundary coverages were created and used in spatial overlay operations to extract digital data for each subbasin from the set of data coverages created specifically for the Sandy and Clackamas river basins.

Hydrography Coverage

The hydrography data layer contains all of the major river tributaries within the Sandy and Clackamas river basins at a 1:100,000 map scale.

The basin boundary of the Sandy and Clackamas drainage basin coverage was used to extract the hydrography data from the Pacific Northwest (PNW) River-Reach files. The PNW River-Reach files contain centerlines that represent stream reaches and connections through reservoirs and braided streams. Each centerline arc is part of a topologically-linked network that allows for upstream and downstream routing processes such as the routing from a headwater stream to the basin outlet. In addition, each arc has attribute information such as stream name, stream reach, and various descriptor codes that are stored in a relational data base file. This information is useful for identifying stream reaches, determining stream lengths and providing descriptor codes that allow for upstream and downstream routing processes.

Precipitation Coverage

Mean monthly and average annual precipitation were collected from National Oceanic and Atmospheric Administration (NOAA) sites and SCS SNOTEL stations throughout Oregon and from nearby locations in adjacent states. Approximately 380 stations make up the data collection network. Point measurements of monthly and average annual precipitation collected during the period of 1961-1990 were evenly spatially distributed across a regular grid using the PRISM model (Daly and Neilson, 1992; Daly and others, 1994). PRISM (Precipitation-elevation Regressions on Independent Slopes Model) is an analytical model, which uses physical and statistical concepts to determine the spatial characteristics of orographic precipitation. PRISM was used to estimate average precipitation values for each 8x8 kilometer grid cell within a grid coverage. A digital isohyetal map of normal

annual precipitation for the State of Oregon was created from the PRISM grid. The basin boundary of the Sandy and Clackamas drainage basin coverage was used to obtain precipitation data from the existing state-wide contour coverage, developed and obtained through Oregon Climate Service (Taylor, 1993). The precipitation contour coverage was used to create the map displayed in Figure 3 which shows the spatial distribution of average annual precipitation in the Sandy and Clackamas river basins.

Land Use Coverage

A 1:250,000 scale digital land use map for the State of Oregon was retrieved from the USGS National Mapping Division's digital database. Detailed information on the development of the land use digital data and the Geographic Information Retrieval and Analysis System (GIRAS) used to store the data is outlined in the USGS digital cartographic data standards circular (Fegeas and others, 1983). Each polygon within the data coverage was coded according to Anderson and others' (1976) Level I land use classification scheme outlined in Table III.

Elevation Coverage

A data layer containing elevation contours was created from a digital elevation model (DEM). A DEM is a data file that includes a set of regularly spaced x,y, and z coordinates that represent a topographic surface, where x and y coordinates represent location and z values represent surface elevation. Two 1:250,000 scale digital elevation models were obtained and merged together to create a single data layer containing the Sandy and Clackamas basins. The DEMs were read into a two dimensional array using ARC/INFO and converted to an ARC grid file. Before merging the coverages

together, each was projected from a geographic projection to a transverse mercator projection. Once a single coverage was created, the Sandy and Clackamas basin boundary coverage was used to extract and create an elevation data layer containing hypsographic contour lines.

TABLE III.
ANDERSON AND OTHERS, (1976) LEVEL I LAND USE
CLASSIFICATION SCHEME.

LEVEL I CODE	CLASSIFICATION
10	URBAN OR BUILT-UP LAND
20	AGRICULTURAL LAND
30	RANGELAND
40	FOREST LAND
50	WATER
60	WETLAND
70	BARREN LAND
80	TUNDRA
90	PERENNIAL SNOW OR ICE

Soils Coverage

Hydrologic characteristics of different soil groups, such as soil permeability and water infiltration rates, are often used to estimate runoff potential. A data layer containing information on the hydrologic characteristics of different soil groups was developed for the Sandy and Clackamas river basins to determine soil permeability rates.

The State Soil Geographic data base (STATSGO) for the State of Oregon was obtained from the SCS. This digital data base contains generalized soil

map units, which were digitized from the State of Oregon General Soils Map. These units were manually compiled from USGS 1:250,000-scale base maps. Each soil map unit contains detailed information on the physical and chemical properties of the soils within that unit. This information is stored in attribute data files linked to the digital soils coverage.

The STATSGO data base was intended to replace the existing hard copy general soil map record with a more consistent digital record containing detailed soil information. STATSGO was developed to be used with a geographical information system —enabling users to store, retrieve, analyze and display soil data more efficiently and effectively.

A soil coverage was created using the basin boundary of the Sandy and Clackamas Basin coverage to extract soil information from the state STATSGO soils coverage. The resulting soil coverage was used to determine an index of soil permeability based on water infiltration rates ranked according to codes A,B, C, and D. Code A represents high soil permeability, and D represents low soil permeability. Codes A through D are used to estimate runoff from precipitation where soils are not protected by a vegetated cover. Soils are assigned to one of these four codes based on the intake of water when soils are thoroughly wetted by precipitation from long duration storms. The soil coverage developed for the Sandy and Clackamas basins indicated that the majority of the study region consisted of soils having moderate infiltration rates and hence assigned a rating code of B.

Water Body Coverages

Two separate coverages containing lakes, ponds, reservoirs and wetland areas were used to compute water storage parameter values.

A 1:100,000 scale digital data layer was obtained from the Pacific Northwest River Reach Banks file, which provided a generalized representation of water storage areas in the Sandy and Clackamas basins. Water storage features, such as flood plain areas, snow fields, and glacial areas, were not represented in this data layer.

Digital wetland inventory information, at a 1:24,000 scale, was retrieved through the Internet from the U.S. Fish and Wildlife Service. This wetland information, previously digitized from 7.5 minute quadrangles, provided a more detailed representation of water storage areas in the Sandy and Clackamas basins than the more generalized 1:100,000 water storage data layer obtained from the Pacific Northwest River Reach Banks file.

COMPUTATION OF BASIN CHARACTERISTIC PARAMETERS USING GIS TECHNIQUES

Six separate GIS computer programs were written using Arc Macro Language (AML). These programs were used to compute all but 4 of the basin characteristic parameters shown in Figure 5. Drainage area, main channel slope, area above 5,000 feet, and a soil permeability index were computed without the use of an AML. The GIS programs were designed to organize and group ARC/INFO commands into a series of instructions, which the computer uses to perform multiple geo-processing operations (such as repetitive spatial overlay procedures among GIS data layers and iterative computations of area and length values). Use of these programs significantly reduced the time and effort required to compute basin characteristic measurements. The six GIS programs, (*ARCSUM.AML*, *MCLLENGTH.AML*, *ELEV.AML*, *CALCLU.AML*, *CALCSTOR.AML* and

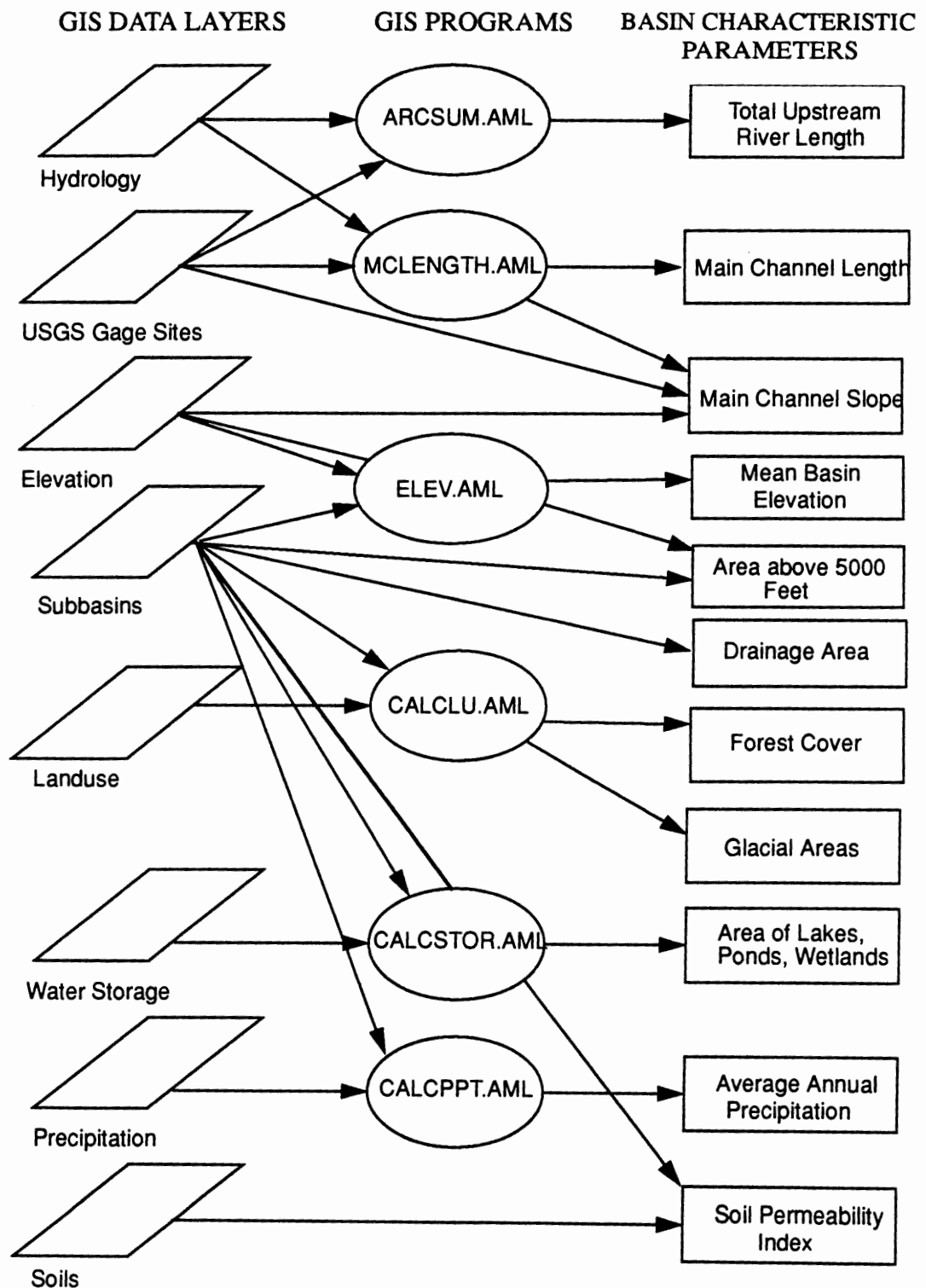


Figure 5. Flow diagram showing the GIS data layers and programs used to compute basin characteristic parameters used in the regression analysis.

CALCPPT.AML) are presented in Appendix E at the end of the report.

Below is a brief description of the 12 climatic and physical basin characteristic parameters computed using GIS automated and semi-automated techniques. Drainage basin characteristic values computed for the 19 subbasins are presented in Appendix C.

Drainage Area

Drainage Area (DA) in square miles, is equal to the total contributing drainage area located upstream from a selected stream site. Drainage basin boundaries were manually drawn on 1:24,000 USGS topographic and input into a GIS digital data layer as polygon features using digitizing techniques. Once in digital format, ARC/INFO utilities automatically compute area values for each drainage area and stores these values in a related INFO file referred to as a polygon attribute table.

Average Annual Precipitation

Average Annual Precipitation (AP) values, in inches, were computed as a weighted average of the total drainage area. Using ARC Macro Language, a GIS computer program (*CALCPPT.AML*) was developed to compute average annual precipitation values for individual subbasin areas. The *CALCPPT.AML* program contains a series of ARC/INFO commands which coordinate GIS spatial overlay operations between the individual *USGS Stream Gage Site* coverages and the *Precipitation* coverage containing average annual precipitation values in inches. The program spatially joins a single *USGS Stream Gage Site* coverage with the *Precipitation* coverage and computes a weighted average annual precipitation value (in inches) for each subbasin area. These values were stored in a related INFO file.

Water Storage

Water Storage (ST) values are expressed as a percentage of the total drainage area and recorded in decimal percent. The GIS program *CALCSTOR.AML* contains a series of ARC/INFO commands used to spatially join the individual *USGS Stream Gage Site* coverages with the 1:100,000 Pacific Northwest River Reach *Water Storage* coverage and again with the 1:24,000 U.S. Fish and Wildlife Service digital wetland inventory *Water Storage* coverage. The percent of the total subbasin area occupied by lakes, ponds, wetlands and reservoirs were computed for each subbasin area at the 1:24,000 and 1:100,000 map scales.

Forest Cover and Glacial Area

Forest Cover (FC) and Glacial Area (GA) values are expressed as a percentage (in decimal percent) of the total drainage area. The 19 individual *USGS Stream Gage Site* coverages were spatially joined with the 1:250,000 scale *Land Use* coverage using the *CALCLU.AML* program to create 19 temporary data coverages. Each temporary data coverage contained area values (in square miles) for each of the Anderson's Level I land use types as well as total area values for each subbasin. Anderson's Level I land use codes were used to identify regions in each subbasin occupied by forest and glacial cover. Forested areas were totalled for each subbasin and divided by the total subbasin drainage area to compute percent forest cover in square miles for each subbasin. The same process was used to compute percent glacier cover values. Percent forest and glacial cover values were stored in a related INFO file.

Main Channel Length

Main Channel Length (MCL) values, in miles, were determined by selecting the longest stream channel above a gaged site. *MCLLENGTH.AML* is a recursive program that uses a series of ARC/INFO commands (similar to those in *ARCSUM.AML*) to compute channel lengths for all contributing stream tributaries located above a gaged site and uses a comparative program loop to identify the longest channel length. Arc attribute codes in the Pacific Northwest River Reach *Stream Network* coverage are used to identify the headwater stream arc. The program routes down the topologically linked stream segment to the end of the arc located just upstream of the gaged site identified using the *USGS Stream Gage Site* coverage. The length of the stream channel from the headwaters to the gaged site is computed and stored in a temporary variable. The program performs the routing process again and computes a new stream channel length. The newly computed channel length is compared to the existing stream channel length stored in the temporary variable. The larger of the two channel lengths is retained and stored as a new temporary variable. The routing routine and computational process repeat until only the largest stream channel length value remains. The program stores the main channel length values for each subbasin in a related INFO file.

Main Channel Slope

Main Channel Slope (MCS) values, in feet per mile, were determined from elevations taken at points near the stream headwaters and near the gaging station using the 1:100,000 scale *Stream Network* coverage and the *USGS Stream Gage Site* coverage. Main channel slope values were determined

by computing the difference in elevation from the headwaters to the drainage basin outlet (USGS gage site) and dividing by the length of the main channel (in miles) computed from the GIS-based computer program *MCLLENGTH.AML*. Main channel slope values were computed using the mathematical equation

$$MCS = \frac{(E_{sh} - E_{bo})}{MCL}$$

where MCS = Main Channel Slope

E_{sh} = Elevation at the stream headwaters (in feet)

E_{bo} = Elevation at the basin outlet (in feet) and

MCL = Main Channel Length (in miles).

It should be noted that there are no universally accepted or standard methods for determining main channel slope using manual or automated techniques. The method described above simply provides a generalized indication of the overall slope of the main stream channel and does not account for abrupt changes in channel slope which is typical of high gradient mountainous streams that flow onto gently sloping low land areas such as those found in the Sandy and Clackamas basins.

Lag-Time Ratio

The lag-time ratio (LT), is a function of main channel length divided by the square root of the main channel slope. As described in Dunne and Leopold (p. 325, 1978), lag-time is computed by the mathematical equation

$$LT = \frac{MCL}{MCS}$$

where LT = Lag-Time

MCL = Main Channel Length (in miles) and

MCS = Main Channel Slope (in feet/mile).

Total Stream Length

Total Stream Length (TSL), in miles, is the total length of all contributing stream tributaries above a gaged site. The GIS program *ARCSUM.AML*, was used to compute total stream length using the *USGS Stream Gage Site* coverage and the 1:100,000 scale Pacific Northwest River Reach *Stream Network* coverage. The *Stream Network* coverage consists of a topologically linked network data layer which allows users to perform directional routing processes such as routing downstream or upstream along a series of connected arcs representing stream reaches. Each arc segment within the *Stream Network* coverage contains important attribute information about each individual stream reach such as stream name, length, river mile and direction of flow. *ARCSUM.AML* contains a series of ARC/INFO commands which are used to identify headwater stream reaches by selecting arc segments that have been manually coded in the attribute file as headwater arcs. Once identified, the program routes in a downstream direction from the headwater arc to the end of the arc located just upstream of the gaged site or selected stream location. Once the routing process is complete, the program computes the length of the stream channel from the headwaters to the gaged site and stores this information in a related INFO file.

Area Above 5000 Feet

Area above 5000 Feet (AF), in square miles, is expressed as a percentage (decimal percent) of the total drainage area. The individual *USGS Stream Gage Site* coverages were spatially joined with the 1:250,000 scale *Elevation* contour coverage containing elevation values in feet. The percent area located above the 5000 foot contour interval was computed for each subbasin area and stored in a related INFO file.

Mean Basin Elevation

Mean Basin Elevation (E), in feet above mean sea level, computed using the GIS-based computer program, *ELEV.AML* to compute mean basin elevation values for each subbasin. The *ELEV.AML* is used to perform a spatial union between a grid representation of the 1:250,000 scale digital elevation model and a grid representation of each subbasin. Mean, minimum and maximum basin elevation values are computed for each subbasin on a cell by cell basis.

Soil Index

Soil codes were determined for the study area by spatially joining the *Sandy and Clackamas Drainage Basin* coverage with the 1:250,000 scale Statewide Statsgo Digital *Soil* coverage to create a new *Soil* coverage for the Sandy and Clackamas river basins. Statsgo soil code information contained in the *Soil* coverage was used to determine soil permeability characteristics within the study area based on the hydrologic soil groups A through D defined by the SCS (Gerig, 1985, Green, 1983). An examination of the digital soil data for the 19 subbasins indicated little variation in soil permeability characteristics over the entire study area. More than 90 percent of the study area was

classified as soil group "B". Based on the SCS classification scheme, soil permeability within the Sandy and Clackamas basins is moderate with infiltration rates ranging from 0.03 to 0.15 inches/hour (Gerig, 1985, Green, 1983).

DEVELOPMENT OF MULTIPLE REGRESSION EQUATIONS

Multiple linear-regression techniques were used to relate flood magnitudes having recurrence intervals of 2, 5, 10, 25, 50 and 100-years to several climatic and physical basin characteristics using the mathematical equation

$$Q_T = K C_1^a C_2^b C_3^c \dots C_n^z$$

in which Q_T is the *response variable* representing estimated flood discharge in cubic feet per second, for a selected T -year recurrence interval; K is a regression constant; C_1 , C_2 , C_3 , and C_n are the *explanatory variables* representing various basin characteristics; and a , b , c , and z are regression coefficients.

The response variable (peak discharge) is assumed to be a linear function of one or more of the explanatory variables (basin characteristics). Flood frequency analyses (Benson, 1964) have shown peak discharge to be linearly related to most climatic and physical basin characteristic variables if the logarithms of each are used. Logarithmic transformations (to the base 10) for both the response variable and the explanatory variables were performed to obtain a linear relation between flood discharges and the various basin characteristics and to establish equal variance of the residuals about the regression line. The Log-transformed regression equation is algebraically equivalent to

$$\text{Log}_{10}Q_T = \text{Log}_{10}K + a \text{Log}_{10}C_1 + b\text{Log}_{10}C_2 + c\text{Log}_{10}C_3 \dots + z\text{Log}_{10}C_n$$

as outlined by Riggs (1968).

Multiple regression techniques were used to develop prediction equations for estimating individual values of flood magnitudes on the basis of one or more significant explanatory basin characteristic variables such as drainage area size, average annual precipitation, and percent forest cover.

Development of multiple regression equations involved several steps. The first was to address the common problem of multi-collinearity by testing for interdependence among the basin characteristic variables using a Pearson (or product moment) cross correlation matrix. Variables showing a high degree of interdependence or strong correlation were not included in the same regression equation.

The second step was to develop several regression equations using different subsets of explanatory basin characteristic variables and assessing the validity of each equation based on statistical methods and basic hydrologic principles.

The final step was to explore the use of surrogate variables. This was to determine whether or not two highly correlated basin characteristic variables could replace one another in a regression model and achieve similar prediction results. In situations where it is too costly or difficult to obtain a particular basin characteristic (such as drainage area), it would be highly beneficial if a surrogate variable (in this case total upstream channel length) could be used in lieu of the missing variable and still be able to provide accurate flood discharge information.

CROSS CORRELATION OF EXPLANATORY VARIABLES

One of the most serious problems encountered in multiple regression analysis is multi-collinearity (Maidement, 1993, Hesal and Hirsch, 1992,

Ott, 1988). This condition occurs when one (or more) explanatory variables is highly correlated with another explanatory variable. The use of highly correlated variables can lead to highly unstable and unrealistic values for the regression coefficients thereby making it difficult to interpret the significance and effectiveness of the individual explanatory variables used in the regression equation.

While it is most desirable to select a set of explanatory variables that are actually independent of each other, this is not always possible in hydrologic analyses since most natural topographic and climatic variables will exhibit some degree of interdependence. For example, factors such as drainage area, main channel length, and total upstream channel length have high degrees of interdependence since main channel length and total upstream channel will increase as drainage areas increase in size.

A computerized procedure for computing Pearson (or product moment) correlation coefficients was performed to test interdependence among twelve selected explanatory variables. This procedure computes a simple matrix showing the degree of correlation among the twelve explanatory basin characteristic variables computed for the 19 subbasins within the Sandy and Clackamas river basins. The correlation coefficients measure the strength of the linear association between two continuous variables. A computed value of 1.0 indicates a perfect correlation, zero indicates complete independence and -1.0 indicates a perfect inverse correlation. The Pearson correlation matrix showing the correlation coefficients for the twelve explanatory variables is displayed in Table IV. Basin characteristic variables, computed for the Sandy and Clackamas river basins, showing a high degree of correlation (>0.8) were not used in the same regression

TABLE IV.

PEARSON PRODUCT - MOMENT CORRELATION COEFFICIENT MATRIX OF EXPLANATORY BASIN
CHARACTERISTIC VARIABLES FOR THE SANDY AND CLACKAMAS RIVER BASINS.

	D-AREA ¹	MCSLOPE ²	MCLENGTH ³	UPLENGTH ⁴	PRECIP ⁵	STORAGE A ⁶	FOREST ⁷	ELEV ⁸	GLACIER ⁹	GT5000 ¹⁰	STORAGE B ¹¹	LAGTIME ¹²
D-AREA	1.00000											
MCSLOPE	-0.57168	1.00000										
MCLENGTH	0.95616	-0.65783	1.00000									
UPLENGTH	0.99853	-0.56039	0.95842	1.00000								
PRECIP	-0.34908	0.64317	-0.38835	-0.32271	1.00000							
STORAGE A	-0.07071	-0.42880	-0.04431	-0.09175	-0.42032	1.00000						
FOREST	-0.02874	-0.53062	0.02049	-0.03854	-0.05976	0.18163	1.00000					
ELEV	-0.25055	0.50771	-0.28377	-0.24578	-0.13020	-0.13621	-0.63859	1.00000				
GLACIER	-0.25486	0.75886	-0.29601	-0.24113	0.27127	-0.33908	-0.86937	0.79522	1.00000			
GT5000	0.24373	-0.73815	0.28948	0.22867	-0.25240	0.36223	0.86535	-0.83493	-0.98856	1.00000		
STORAGE B	-0.10914	-0.19095	-0.13384	-0.13078	-0.65333	0.55344	-0.05434	0.28847	-0.06016	0.06216	1.00000	
LAGTIME	0.98255	-0.61320	0.98475	0.98102	-0.37839	-0.04627	-0.02456	-0.28801	-0.28047	0.27513	-0.10635	1.00000

A value of 1.00000 indicates a perfect correlation. Shaded values indicate a high degree of correlation with values greater than +0.80 and less than -0.80.

¹D-AREA, drainage area, in square miles,

²MCSLOPE, main channel slope, in feet per mile,

³MCLENGTH, main channel length, in miles,

⁴UPLENGTH, total upstream channel length, in miles,

⁵PRECIP, average annual precipitation, in inches,

⁶STORAGE A, area of lakes, ponds, wetlands and reservoirs, in decimal percent plus 0.01. area values were computed from the Pacific Northwest River Reach hydrography data layer (1:100,000 map scale).

⁷FOREST, forest cover, in decimal percent,

⁸MEAN-ELEV, mean basin elevation, in feet,

⁹GLACIER, glacier areas, in decimal percent,

¹⁰GT5000, area above 5000 feet, 1 minus the decimal percent,

¹¹STORAGE B, area of lakes, ponds, wetlands, in decimal percent. Area values were computed from the U.S. Fish and Wildlife Service digital wetland inventory data layer (1:24,000 map scale),

¹²LAGTIME, lag time, main channel length divided by the square root of the main channel slope.

equation. Variables showing a high degree of correlation (>0.8) were: (a) drainage area, main channel length, lag-time, and total upstream length, (b) forest cover, glacier areas, and areas above 5000 feet elevation and (c) mean basin elevation and areas above 5000 feet.

Main channel length and total upstream channel length show a strong correlation with total drainage area with correlation coefficients of 0.95616 for main channel length and 0.99853 for total upstream stream length. Intuitively, it is easy to recognize that total channel lengths will increase in size as drainage basin areas become increasingly larger. Main channel length and total upstream channel length are also highly correlated with a correlation coefficient of 0.95842. Lag-time ratio and drainage area show a strong correlation since an increase in drainage area often results in an increase in lag-time since runoff must travel across increasingly larger expanses of land surface. Lag-time is a function of channel length and slope and is expressed as a lag-time ratio which is the main channel length divided by the square root of the main channel slope.

Values computed for percent glacial area and area above 5000 feet are negatively correlated as indicated by the correlation coefficient of -0.98856. Both variables are similar indicators of higher elevations where precipitation is mainly in the form of snow. These factors can be useful for explaining flood peaks by indicating snow melt events or accounting for a reduction in runoff by functioning as storage areas. These two parameters indicate a high degree of correlation with forest cover, with correlation coefficients of -0.86937 for glacier areas and 0.86535 for areas above 5000 feet. Mean elevation is also correlated with these two parameters with correlation coefficients of -0.83493 for areas above 5000 feet and 0.79522 for glacier areas.

MODELS INVESTIGATED

Regression models consisting of different subsets of explanatory basin characteristic variables were investigated. The validity of each model was assessed based on statistical methods and basic hydrologic principles. All 12 of the selected basin characteristic variables can be significantly related to peak flow however, only six variables were used in the final regression analysis. Those basin parameters not retained in the analysis were either highly cross correlated to the selected six variables or found to be less significant in a preliminary analysis.

The six explanatory basin characteristic variables used in the regression analysis were: (1) drainage area, (2) average annual precipitation, (3) area above 5000 feet, (4) main channel slope, and (5) water storage at the 1:100,000 map scale and (6) water storage at the 1:24,000 map scale. Within the study area, the range of values for these variables are listed below in Table V.

TABLE V.

BASIN CHARACTERISTIC PARAMETERS AND RANGE OF VALUES
FOR THE SANDY AND CLACKAMAS RIVER BASINS, OREGON.

BASIN CHARACTERISTIC PARAMETERS	RANGE OF VALUES
Drainage Area	3.78 - 933.94 square miles
Average Annual Precipitation	49.6 - 107.6 inches
Area Above 5000 Feet	0 - 40.33 percent of total basin area
Main Channel Slope	60 - 640 feet per mile
Water Storage A (1:100,000 map scale)	0 - 1.64 percent of total basin area
Water Storage B (1:24,000 map scale)	0.09 - 5.9 percent of total basin area

Flood frequency data and basin characteristic values were input into a statistical computer software system called Statit (Statware Inc., 1992). A logarithmic transformation of the data was performed to establish a linear relation between peak discharge and the selected basin characteristics. Computerized computations were performed using a multiple regression "subset" approach which examines all possible subsets of explanatory variables that best predict the response variable. The Statit (Statware Inc., 1992) regression program, "ALLREG", examines all possible subsets of explanatory variables and identifies the "best one parameter model", the "best two-parameter model", the "best three-parameter model", and so on. From these resulting models, the program then selects the "best overall model" based on the selection criteria of minimizing the Mallows' C_p statistic (Mallows, 1973).

In addition to the Mallows' C_p statistic, the program also computes the Akaike' information criterion (AIC), the adjusted R-square (R^2), the root mean square error (MSE), an analysis of variance table and a table of regression coefficients for each model. The usefulness of each explanatory variable is determined based on its statistical significance and on the percent reduction in the standard error of estimate as each additional variable is added to the model. The benefit of adding additional variables to a regression model is to account for or explain more of the variance of the response variable (Hesel and Hirsch, 1992). The cost, however, is that the degrees of freedom decrease and the width of the confidence intervals increases making it more difficult to establish statistical significance in the hypothesis tests.

Because the adjusted R^2 adjusts for the degrees of freedom in the model and penalizes a model that includes too many slope parameters, it is important to note when the introduction of additional explanatory variables shows a decrease or an increase in the adjusted R^2 (the coefficient of determination adjusted for the number of explanatory variables used in the regression).

Two-Parameter Model

The results of the regression analysis indicated that the the best two-parameter model for defining peak discharge in the Sandy and Clackamas basins included drainage area and average annual precipitation. Results of previous flood frequency analyses for this region have consistently shown drainage area to be the most statistically significant explanatory variable affecting peak discharge. After drainage area, a measure of precipitation was found to be the next most significant variable for defining flood magnitudes in this region.

A plot of the residuals (Figure 6) using the two-parameter model show outlier values for several of the gaging stations, some of which can be explained. Station 14131000 (Little Zigzag at Twin Bridges) and 14131400 (Zigzag River near Rhododendron) are outliers. These drainage basins (Figure 4) are located in the higher elevations on the western slopes of Mt. Hood. Approximately 9 percent of the drainage basin above station 14131000 (Little Zigzag at Twin Bridges) and 5 percent above station 14131400 (Zigzag River near Rhododendron), are covered by snow and glacial fields. Approximately 1/3 to 3/4 of the total annual precipitation falls as snow in these basins (State of Oregon Water Resources Board, 1965). For most of the year, precipitation is

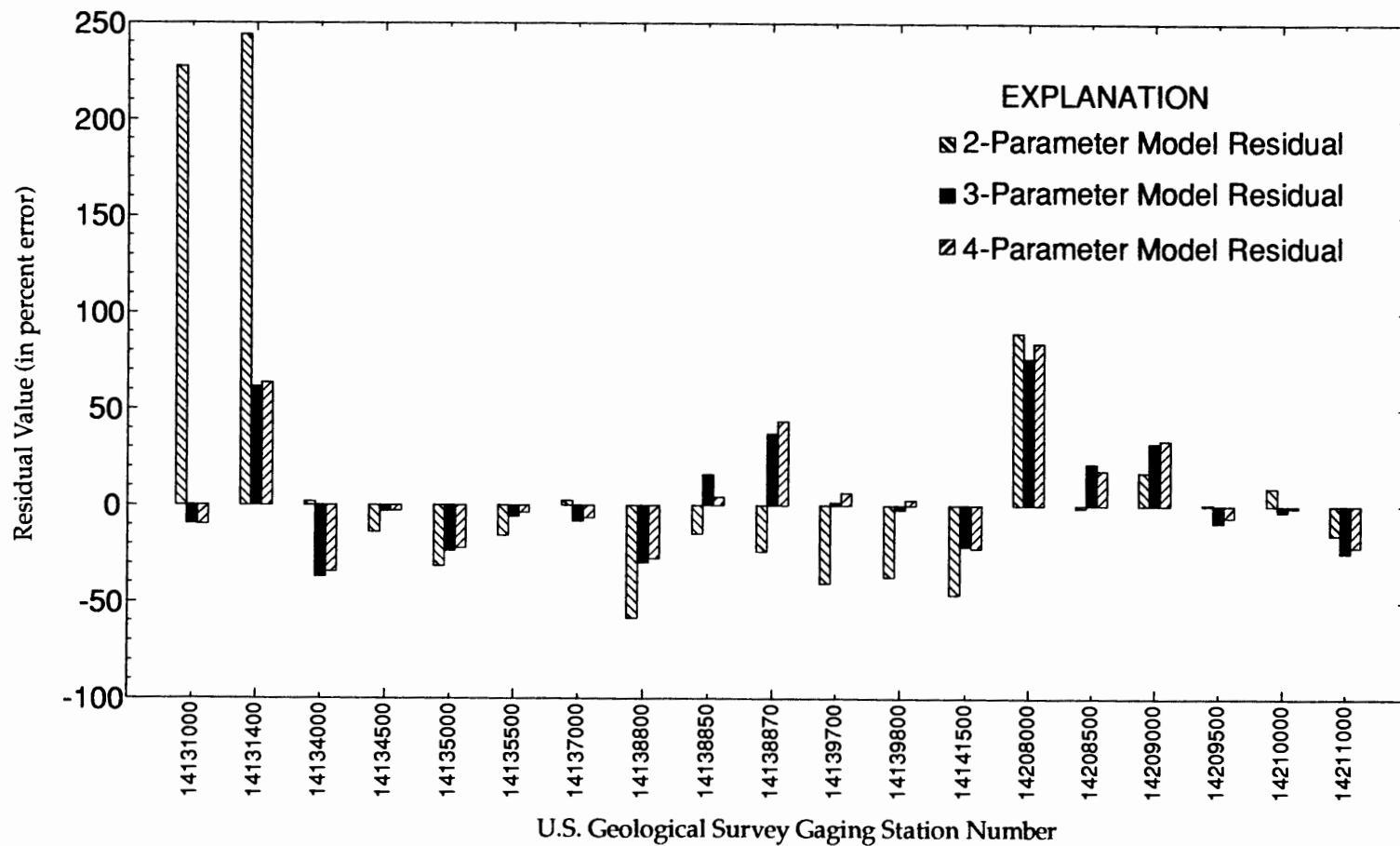


Figure 6. Plot of the residuals of predicted flood magnitudes for the 2-year recurrence interval using the 2, 3, and 4-parameter regression equations.

retained in storage and does not become direct runoff and contribute to peak flow conditions until the melt season begins. As a result, drainage area and average annual precipitation alone do not sufficiently explain the variations in peak discharge occurring in the Sandy and Clackamas basins especially for sites located in the higher elevations where precipitation occurs mainly in the form of snow.

Three-Parameter Model

The regression relation of peak discharge to basin characteristics was significantly improved when the variable representing percent area above 5000 feet was added to the two-parameter model containing drainage area and average annual precipitation. Its inclusion significantly improved the regression relation and brought in the two outliers for the stations, 14131000 (Little Zigzag at Twin Bridges) and 14131400 (Zigzag River near Rhododendron). As shown in Table VI, the adjusted R^2 value increased from 0.878 in the two-parameter model to 0.967 for the three-parameter model indicating a significant improvement by the inclusion of the third explanatory variable.

The reduction in the mean square error from 0.0627 to 0.0172 and the Mallows' C_p statistic from 38.86 to 2.34 also indicated a significant improvement in the regression relation. The three-parameter model improved outlier values for stations 14131000 (Little Zigzag at Twin Bridges) and 14131400 (Zigzag River near Rhododendron) as shown in Figure 6.

TABLE VI.
COMPARISON OF 2-YEAR RECURRENCE INTERVAL REGRESSION
EQUATIONS USING DIFFERENT SUB-SETS OF EXPLANATORY
VARIABLES.

Explanatory variables in model	Explanatory Variables	Adjusted R ²	Mean Square Error ¹	Mallow's Cp	Percent Standard Error Estimate
2	DA, AP	0.8784	0.0627	38.86	60.8
3	DA, AP, AF	0.9667	0.0172	2.34	30.7
4	DA, AP, AF, STA	0.9649	0.0181	4.11	31.5
4	DA, AP, AF, STB	0.9747	0.0168	4.03	30.3
4	DA, AP, AF, MCS	0.9645	0.0183	4.23	31.6

¹ Mean square error of log units

DA = Drainage Area
 AP = Average Annual Precipitation
 AF = Percent Area above 5000 Feet
 STA = Water Storage A (1:100,000 map scale)
 STB = Water Storage B (1:24,000 map scale)
 MCS = Main Channel Slope

Four-Parameter Models

Water storage was reported as being one of the most important explanatory variables for defining flood magnitudes for the High Cascade region (Harris and others, 1979). It was therefore hypothesized that the introduction of a water storage parameter to the three-parameter model would likely improve the regression relation and explain more of the variation in peak discharge characteristics in the Sandy and Clackamas basins. Surface area values for lakes, ponds, wetlands and reservoirs were computed from a generalized 1:100,000 scale digital data layer obtained from

the Pacific Northwest River Reach Banks digital data layer. Computed values obtained from this data layer were represented as Water Storage "A" (STA). As shown in Table VI, the addition of water storage (STA) as a fourth variable, did not improve the regression relation but provided satisfactory estimates of predicted peak flow values.

It was believed that better prediction results could be achieved by obtaining a better indication of water storage using a higher resolution digital data. A 1:24,000 scale GIS digital data layer containing detailed information on wetland areas in the Sandy and Clackamas basins was obtained from the U.S. Fish and Wildlife Service. Water storage values were re-computed and added to the three-parameter model as Water Storage "B" (STB).

Results of the regression analysis showed that the inclusion of the new storage data showed a slight improvement in the adjusted R^2 and Mallow's C_p statistic when compared to the four-parameter model containing the more generalized storage information. However, the coefficient of the Water Storage "B" parameter in this four-parameter model was positive indicating that an increase in storage results in an increase in flood magnitudes. Based on hydrologic experience, this model was contrary to known effects of water storage on peak discharge characteristics. For this reason, the four-parameter model with Water Storage "B" was rejected even though it showed an improvement.

The addition of a storage variable did not improve the regression model using either one of the two data layers representing storage areas. The most probable reason for this failure is that neither data layer contained

information on flood-plain storage which is probably the most important storage component in reducing flood peaks (Laenen, 1983).

Best Overall Model

The best regression model is one that will explain as much of the variance of the response variable as possible with the smallest number of explanatory variables. The best overall regression model was determined by evaluating the statistical significance of the selected explanatory variables identified by the computerized computations, by noting the reduction of the standard error as explanatory variables were included or eliminated from the regression model and by minimizing the Mallow's C_p statistic. This was further followed by an assessment of validity based on basic underlying hydrologic principles.

Of the regression models investigated, the three-parameter model was determined the best overall model for explaining variations in peak discharge and for determining flood magnitudes for selected recurrence intervals in the Sandy and Clackamas river basins. The explanatory basin characteristics that best define flood magnitudes were, in order of significance, drainage area (DA), average annual precipitation (AP), and area above 5000 feet (AF). Although previous flood frequency analyses have found water storage to be a significant explanatory variable for defining flood magnitudes for this study region, this parameter could not sufficiently be quantified and was therefore not included in the final model.

The regression equations for the three-parameter model as well as the average absolute percent error, the adjusted R^2 , and the percent standard error for the 2, 5, 10, 25, 50 and 100-year recurrence intervals are shown in Table VII. An adjusted R^2 value of 0.97 indicates that 97 percent of the

TABLE VII.

THREE PARAMETER REGRESSION EQUATIONS FOR THE SANDY AND CLACKAMAS RIVER BASINS. THE THREE-PARAMETER MODEL PROVIDES THE "BEST OVERALL" MODEL FOR DEFINING FLOOD MAGNITUDES FOR SELECTED RECURRENCE INTERVALS WITHIN THE STUDY REGION.

Exceedance Probability (RI) ¹	EQUATIONS Sandy and Clackamas River Basins (19 Stations)	Average Absolute Percent Error	Adjusted R ²	Percent Standard Error
Q _{0.5} (2)	0.0000635 (DA) ^{0.940} (AP) ^{3.17} (AF) ^{3.71}	22	0.97	31
Q _{0.2} (5)	0.0001788 (DA) ^{0.948} (AP) ^{3.01} (AF) ^{3.47}	25	0.96	33
Q _{0.1} (10)	0.0003281 (DA) ^{0.952} (AP) ^{2.91} (AF) ^{3.32}	27	0.95	37
Q _{0.04} (25)	0.0006591 (DA) ^{0.955} (AP) ^{2.79} (AF) ^{3.14}	29	0.95	39
Q _{0.02} (50)	0.0010627 (DA) ^{0.957} (AP) ^{2.71} (AF) ^{3.01}	31	0.94	41
Q _{0.01} (100)	0.0016638 (DA) ^{0.958} (AP) ^{2.63} (AF) ^{2.89}	33	0.94	43

General form of equation $Q_T = K(DA)^a(AP)^b(AF)^c$ where

Q_T = discharge for selected exceedance probability

K = regression constant

DA = drainage area

AP = average annual precipitation

AF = area greater than 5000 feet, expressed as 1 - (GT, in decimal percent)

¹ Numbers in parentheses refer to recurrence intervals in years.

variation in flood discharge can be explained by the contributing drainage area, average annual precipitation, and percent of the basin area located above 5000 feet. The standard error of estimate provides a measure of the reliability of the regression equation and represents the standard deviation

of the distribution of the residuals about the regression line. It was computed by transforming the root mean square error from log units back to normal using methods described by Riggs (1968). The standard error of estimate for the three-parameter model ranged from 31 to 43 percent. These values indicate the percent error associated with the predicted values for peak discharge.

Using the three-parameter model, flood magnitudes for 2-, 5-, 10-, 25-, 50-, and 100-year recurrence intervals can be computed for any stream location, gaged or ungaged, within the study region by inputting the appropriate basin characteristic values into the established regression equations.

For example, flood magnitudes can be estimated at a proposed bridge site within the study region by obtaining values for contributing drainage area, average annual precipitation and percent of the drainage area above 5000 feet and inputting these values into the three parameter regression equations. Basin parameter values can be rapidly developed and computed using GIS techniques and AML algorithms provided in this study.

It is important to note that the equations developed in this study are valid for the Sandy and Clackamas river basins and are only applicable to stream sites that represent natural streamflow conditions and have no (or minimal) upstream diversions or regulation. The flood equations are limited to the range of parameter values used in the analysis and conditions sampled by the defining data. Extrapolation beyond these characteristic limits could produce erroneous results.

FLOOD PREDICTION ACCURACY - GIS VERSES MANUAL METHODS

Flood magnitudes for the 2-year recurrence interval were computed for 12 gaging stations using regression equations developed by Harris and others (1979) in a previous flood frequency analysis for western Oregon using manual methods. Flood magnitudes for the 2-year recurrence interval were also computed for the same 12 gaging stations using the three parameter regression equation developed in this study using GIS techniques. These computed flood magnitudes were then compared to flood magnitudes computed using the Log Pearson Type III flood frequency distribution to compute percent error values for the 12 gaging stations and average absolute percent error. The purpose of this was to compare the level of accuracy for flood predictions using GIS verses manual methods. Flood magnitudes (in cubic feet per second) and percent error for the 12 gaging stations are shown in Table VIII.

The average absolute percent error computed for the three parameter model developed using GIS techniques was 23 percent indicating a slight improvement over Harris and others' models which had an average absolute percent error of 28 percent. The three-parameter model significantly improved the outlier value for Station 14208000 (Clackamas at Big Bottom) by reducing the residual error from +113 to +76 and brought in the outlier values for stations 14138850 (Bull Run River near Multnomah Falls) and 14141500 (Little Sandy River near Bull Run).

The standard error of estimate is significantly improved when comparing the three-parameter model developed in this study to the four-parameter model developed by Harris and others (1979) for the High Cascade region. Standard error of estimates for the Sandy and Clackamas

TABLE VIII.
COMPARISON OF PREDICTED 2-YEAR FLOOD MAGNITUDES
COMPUTED USING REGRESSION EQUATIONS DEVELOPED
USING GIS AND MANUAL METHODS

LOG PEARSON TYPE III		GIS METHODS		MANUAL METHODS	
Computed peak discharge was determined using the Log Pearson Type III flood frequency distribution		3 Parameter Model for Sandy and Clackamas Region $0.0000635(DA)^{0.940}(AP)^{3.17}(AF)^{3.71}$		Harris and others (1979) Models for Willamette Region $8.70(A)^{0.87}(T)^{1.7}$ High Cascade Region $4.75(A)^{0.90}(ST+1)^{-0.62}(101-F)^{0.11}(I)^{1.17}$	
STATION NUMBER	COMPUTED PEAK DISCHARGE (ft ³ /s)	PREDICTED PEAK DISCHARGE (ft ³ /s)	PERCENT ERROR (pred-comp/comp) x 100	PREDICTED PEAK DISCHARGE (ft ³ /s)	PERCENT ERROR (pred-comp/comp) x 100
14134000	283	177	-37	230 ^b	-19
14134500	1,361	1,320	-3	1,057 ^b	-22
14135000	5,359	4,101	-23	5,117 ^a	-5
14137000	14,500	13,300	-8	12,340 ^a	-15
14138800	1,096	771	-30	708 ^a	-35
14138850	5,790	6,711	+16	3,299 ^a	-43
14141500	2,157	1,691	-22	1,104 ^a	-49
14208000	3,008	5,298	+76	6,403 ^a	+113
14208500	504	613	+22	503 ^b	0
14209000	1,659	2,180	+32	2,107 ^b	+27
14209500	17,141	15,935	-7	18,316 ^a	+7
14210000	24,636	23,847	-3	24,558 ^a	0
		Average Absolute Percent Error = 23		Average Absolute Percent Error = 28	

^a Values computed using regression equation for the High Cascade region.

^b Values computed using the equation for the Willamette region

river basins range from 31 to 43 for the three-parameter model compared to 55 to 72 for the High Cascade region which represented a much broader area. For the Willamette region, standard error of estimates range from 33 to 37 using a two-parameter model which are comparable to the three-parameter model.

The results from the regression analysis indicated that the three-parameter model developed using GIS techniques provided better estimates of peak flow than the models developed using manual methods. It is difficult however, to determine whether the improved level of accuracy for flood prediction can be attributed to the use of GIS technology or due to the smaller sample size and smaller study region used in this analysis. For instance, the regression equations developed for the Sandy and Clackamas basins account for variations in peak discharge for a relatively homogeneous climatic region containing a small sample size of 19 gaging stations. Whereas the equations developed by Harris and others (1979) account for variations in peak flow characteristics for larger study regions using 111 gaging stations for the Willamette region and 28 stations for the High Cascade region.

INVESTIGATION OF SURROGATE VARIABLES

It is important to investigate the possibilities of using surrogate variables in instances where it would be difficult to obtain or develop a particular basin characteristic parameter. For example, the delineation of contributing drainage area is essential for all flood frequency regression analyses since it is the most significant explanatory variable in most flood equations. Drainage areas must be hand delineated on topographic maps

regardless of whether these delineations are to be digitized using GIS hardware or traced manually using planimeters. In this application, the use of GIS techniques is no less time consuming or tedious than manual methods. However, there are several advanced GIS-based algorithms that will automatically delineate drainage basin boundaries (Jenson, 1988, 1991, ESRI, 1990) using digital elevation models (DEMs) and other terrain models such as triangular irregular networks (TINs). Unfortunately however, these algorithms have several limitations and until these are remedied, the problem of obtaining drainage area delineations remains. In light of this, a search was made to identify a surrogate variable to take the place of drainage area in the regression equation.

Drainage area is highly correlated with total upstream channel length (which is computed by summing the channel lengths of all contributing tributaries located upstream of a gaged site) with a correlation coefficient of 0.99853 (Table IV). These two variables appear to be linearly related when the logarithms of both are plotted (Figure 7).

Station 14131000 (Little Zigzag at Twin Bridges) and station 14131400 (Zigzag River near Rhododendron) are the two outliers. Drainage basins above these gaged sites are located at higher elevations just below glaciated areas found on the slopes of Mt. Hood. Drainage areas are small for these sites and stream lengths are minimal since much of the upper reaches of the basin are covered by glacial snow fields.

Since total upstream channel length and drainage basin area are highly correlated and linearly related, it was believed that total upstream channel length could replace drainage area in the regression model and by doing so, could produce similar results and provide a similar level of accuracy for

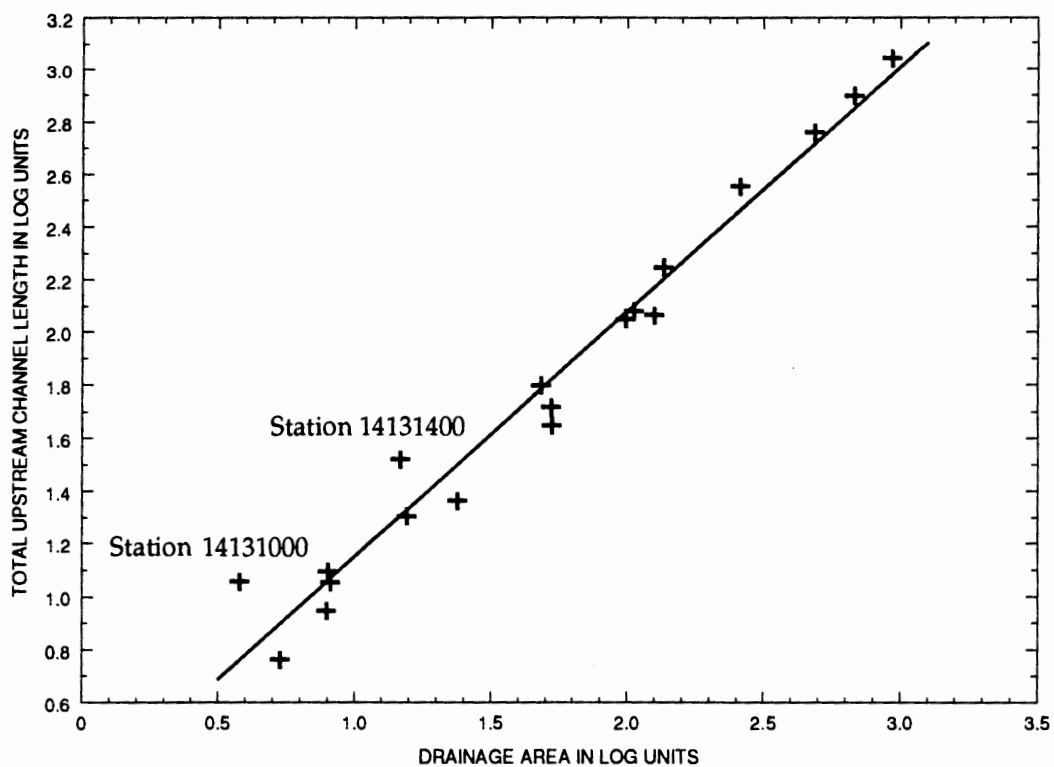


Figure 7. Plot showing the linear relation between the log of Drainage Area and the log of Total Stream Length.

flood prediction. To test this hypothesis, total upstream river length replaced drainage area in the three-parameter model. The results of the regression analysis using total upstream channel length were compared to the results achieved using drainage area (Table IX).

Based on the results of the analysis, a comparable level of accuracy of flood prediction for the Sandy and Clackamas river basins could be achieved using total upstream channel length as a surrogate variable for drainage area. Total upstream river length can easily be obtained from the hydrology

TABLE IX.
COMPARISON OF THREE-PARAMETER MODELS FOR THE 2-YEAR
RECURRENCE INTERVAL. TOTAL UPSTREAM RIVER LENGTH REPLACED
DRAINAGE AREA IN THE SECOND REGRESSION MODEL.

Explanatory variables in model	Explanatory Variables	Adjusted R ²	Mean Square Error	AIC	Mallow's Cp
3	DA, AP, AF	0.9667	0.0172	-73.662	2.34
3	TSL, AP, AF	0.9652	0.0179	-72.869	4.00

DA = Drainage Area
AP = Average Annual Precipitation
AF = Percent Area above 5000 Feet
TSL = Total Stream Length
AIC = Akaike's Information Criterion

data layer using a GIS-based algorithm specifically designed to rapidly and automatically compute these values. Use of this variable in replace of drainage area values is highly desirable in situations where the delineation of drainage area requires extensive time and effort.

DISCUSSION

Results of the regression analysis indicated drainage area, average annual precipitation, and area above 5,000 feet are the most statistically significant basin parameters for predicting flood magnitudes in the Sandy and Clackamas basins. Harris and others' (1979) flood frequency analysis of the Willamette region (which includes more than 75 percent of the Sandy and Clackamas basins) also indicated drainage area and precipitation intensity as being the most significant explanatory variables for determining flood magnitudes in the Willamette region. Standard errors for the

Willamette region ranged from 33 to 37 percent and compared reasonably well with those computed for the Sandy and Clackamas basins which ranged from 22 to 33 percent. The High Cascade region (which includes the higher elevations of the Sandy and Clackamas basins) however, had significantly higher standard error of estimates that ranged from 55 to 72 percent.

The most significant basin parameters included in the regression equations for the High Cascade region were drainage area, area of lakes and ponds, forest cover, and precipitation intensity. These equations lacked indicators of snowpack and snow melt processes which have a profound influence on runoff and peak flow characteristics in these higher elevations. Harris and others (1979) noted that the higher standard errors computed for the High Cascade region could possibly be attributed to the lack of regional snowpack information and better regression results could possibly be achieved by the inclusion of a basin parameter for snowpack. Reliable snowpack information however, was not available during the time of the study.

Based on the regression results, the three parameter regression equations can be used to compute flood magnitudes for select recurrence intervals for ungaged stream locations in the Sandy and Clackamas river basins by computing values for drainage area, average annual precipitation and area above 5000 feet. Improved estimates may possibly be achieved by the addition of a better defined basin parameter for water storage. Neither the 1:100,000 scale digital data layer containing surface area values for lakes, ponds, wetlands and reservoirs nor the 1:24,000 scale digital data layer containing detailed wetland information were useful in the regression relation for defining flood magnitudes.

Obtaining reliable estimates of water storage within a drainage basin continues to be problematic for hydrologists. During floods, water storage in lakes, ponds, wetlands and reservoirs is not the only water storage that occurs in a basin and may not be the most important in flood peak reduction. Lack of data on water storage facilities such as flood plains, stream channel storage, soil and groundwater storage and surface depressions make it difficult to account for the dispersion, retention, and temporary of surface water runoff. Flood plain information, for example, is important for defining flood peaks and was reported as being the most significant variable (next to drainage area) in a flood analysis of urban areas (Laenen, 1980). Flood plain information however, is not available for most rural and undeveloped areas.

The plot in Figure 6 shows high residual values for station 14208000 (Clackamas at Big Bottom). The two, three and four-parameter models did little to improve these residual values. Peak discharge values computed from the three models were significantly higher than actual peak flows determined at the Clackamas at Big Bottom stream gage. The great disparity in peak flow estimates for this basin could possibly be improved by providing an explanatory variable for flood plain areas and a better indication of water storage. None of the models account for the large flood plain area located directly upstream of the Big Bottom gaging station. This flood plain area will significantly reduce flood peaks occurring at this gage site. Since the regression models do not account for flood plain areas, predicted peak flows are significantly higher than actual peak flows. Water storage information provided in the four-parameter models do little to explain what is occurring in these basins. For example, the 1:24,000 data layer containing detailed wetland information was not adequate for

explaining variations in peak flow since many of the wetland areas identified are located a great distance from the stream channels and therefore have little affect on peak flow conditions.

To remedy the situation, it may be possible to explore the use of surrogate indicators of flood plain areas and areas where water may be temporarily stored or retained using GIS techniques. It would be useful to explore the use of GIS to identify flood plain areas and channel storage using high resolution digital data to measure average stream width and length of stream meanders. Another alternative would be to develop a GIS-based computer programs to delineate "flat areas" adjacent to stream channels. This approach however, requires considerable time and effort to develop.

SUMMARY OF THE BENEFITS AND LIMITATIONS OF APPLYING A GIS

This study is an important step in determining the feasibility of applying a GIS in a flood frequency analysis. It provides an opportunity to explore the use of newly developed GIS techniques, such as the use of GIS-based computer programs, and an opportunity to address many of the current limitations and constraints of manual methods encountered in previous flood frequency analyses.

Improved regression results achieved in this flood study can be attributed to the use of GIS by allowing for the exploration and development of new basin parameters previously not considered in multiple regression flood frequency analyses. For example, GIS allowed for the development of basin parameters such as glacial areas, areas above 5000 feet, total upstream channel length and two different resolutions of water storage information. Area above 5000 feet turned out to be the third most significant basin characteristic for determining flood magnitudes in the Sandy and Clackamas basins. Use of this parameter significantly improved the regression results and brought in outliers representing gaged sites located in the higher elevations.

Not only did GIS allow for the exploration and development of new data parameters, it made it possible to change various model parameters during the regression analysis process. For example, water storage information obtained from a generalized 1:100,000 scale digital layer was found to be insignificant in the preliminary regression analysis. To obtain

this information for the Sandy and Clackamas study region using manual methods would require the acquisition of 46 separate 7.5 minute quads containing wetland information and several months of computing area values using either a hand held planimeter or transparent grid overlay techniques.

The most valuable asset of the GIS data base was the ability to rapidly and interactively change various model parameters as necessary by having spatial data readily available. High resolution digital wetland data, (at a 1:24,000 map scale) previously developed by the U.S. Fish and Wildlife Service (1995), were obtained over the Internet. A GIS-based algorithm was developed to automate the computational process of calculating lake, pond, wetland and reservoir values areas wetland. Once computed, these values were brought into the regression model and the regression and the regression analysis were re-run. The whole process took little more than a week to accomplish. By having standardized digital data already available, less time was spent on the development of the data base and more time was spent on the calibration of the model.

An essential benefit of a GIS lies in the use of previously developed data that, once in a standardized digital format, can be easily updated and maintained on a continual basis. The initial cost of developing a large GIS data base is offset by the future convenience and flexibility in modifying and updating information as new data becomes available and as changes in the hydrologic regime occur. Updates of flood frequency analyses can be performed in a relatively short amount of time and with relative ease. By using an established GIS data base, it is no longer necessary to start the flood analysis from scratch. Data simply needs to be added to the established data

base as it is collected. The use of GIS is a vast improvement over the use of hand held planimeters since data is put into digital format where it can be used again in future applications.

Use of a standardized format allows GIS digital data to be shared among different agencies and user groups. In previous flood studies, much of the data was in fixed formats making it difficult if not impossible to be used with other data formats used in other studies. In addition, most conventional data bases are made up of separate files that cannot be spatially overlaid or combined for analyses or used in a relational data base. Each file is independent of other files. The development of new data sets at the onset of each project caused unnecessary duplication of effort and unwisely spent time and money.

Once developed and maintained in a standardized format, the GIS data layers can be easily updated for future use and additional applications that may overlap the study region. Data can be shared by outside agencies and easily transferred through FTP or over the Internet. Updates simply require retrieving the data layer and updating the topology or adding or editing new information in the related attribute file.

The most valuable asset of a GIS was the use of advanced GIS-based programs (AMLs). The compilation and computation of the various hydrologic and climatic components were performed rapidly and efficiently through the use of these programs. The use of these AMLs were extremely useful in situations involving multiple spatial overlay processes and extensive mathematical computations.

There was an initial time investment in developing AML programs and a trial and error process of de-bugging the program errors. Writing these

programs required extensive knowledge of GIS commands and processes and a certain level of understanding of computer programming. The time required to write and modify these programs will depend largely on the level of expertise of the programmer and the complexity of the program at hand.

Geographical information systems offered several specialized spatial analysis capabilities which could be further automated using advanced GIS-based computer algorithms. For example, determining the best method for delineating basin boundaries remains a problem. The manual delineation of the contributing drainage area above a selected stream site from a topographic map is a time consuming and tedious task, especially when the number of sites is large. However, the process of digitizing drainage basin boundaries using GIS hardware and software components is no less tedious and time consuming than using hand-held planimeters to perform the same task. The benefit of using a GIS over manual methods is realized once the data is stored in digital format. Once in place, the data could be manipulated and updated with relative ease while manual method leave little to be gained with respect to future applications.

However, automated techniques for delineating basin boundaries have limitations and are not without serious pitfalls. Problems occur in areas that have minimal topographic relief since the program has difficulty distinguishing basin ridges in regions that are relatively flat. Problems also occur in urbanized or developed areas where artificial structures and diversions alter the landscape and control the flow of water. Pits and depressions in the landscape also cause problems since the program has difficulty interpreting the terrain.

There is a need for a reliable automated approach to delineating basin boundaries. Algorithms have been developed that will automatically delineate drainage basin boundaries using a digital elevation model (DEM) (Jenson, 1991). These, however, have had limited success and often require the use of high resolution DEMs at 1:24,000 scales. DEMs at this resolution are unfortunately not available for most regions and 1:250,000 DEMs are often not sufficient for accurately delineating basin boundaries in many areas.

A serious limitation of this study is that all of the AML programs used in this analysis require a basin boundary to extract the basin parameter information from other data layers. Without the basin boundary, these programs cannot be successfully run. Until an alternative technique is developed, drainage basin delineations must be established by the user.

Admittedly, GIS requires large initial investment in time and effort compared to traditional methods, once made however, data bases are easily updated as changes take place. It is anticipated that the time and effort involved in data preparation and maintenance will be reduced as digital maps become increasingly available.

RECOMMENDATIONS FOR FUTURE STUDIES

Flood prediction and flood hazard research is an evolving science. Improvements in flood prediction will depend on the continued development of new technologies, use of up-to-date and reliable data, and evolution of better methods and indices for describing and defining the physical and climatic factors that influence flood magnitudes.

Based on the results of this study, there is a need for reliable methods for automating the process of drainage area delineations. Work is currently being done to address the limitations encountered in algorithms designed to automatically delineate basin boundaries. Presumably, the success rate of these programs will largely depend on the increase availability of high resolution digital elevation models. In the interim, it may be necessary to explore the use of GIS techniques to semi-automate the process of basin delineations above a stream site. These might include using the GIS programs to buffer around stream channels until these boundary lines coalesce to create a single basin boundary. Another suggestion is to create a generalized basin boundary by connecting the end points (or nodes) of the headwater arcs and expanding this area by a designated factor such as five percent. Lastly, it may be necessary for the user to interactively draw a basin boundary above the stream site of interest using the computer mouse to draw the outline on the computer screen. Much is to be gained by the successful development of a reliable method for automatically obtaining basin boundaries.

The most promising use of GIS in future flood frequency applications is the eventual development of a fully-automated, menu-driven computer interface that will allow both GIS and non-GIS professionals to “point-and-click” on the stream location of interest and automatically obtain flood information for desired recurrence intervals. Development of such a product would be highly beneficial to engineers, water resource managers, planners and other users needing reliable up-to-date flood information. Pull down menus could be developed to make the package user friendly and easy to use. This would require the development of digital data coverages for each of the necessary basin parameters for the entire State of Oregon. Several GIS-based computer programs would need to be developed to fully automate the data extraction procedures. Ideally, the data sets required for this application would be relatively small such that the program could be used on personal computers making it more accessible to a wider user group. The future looks promising for the development of such a package, however, it will require extensive time and effort and the expertise of highly-trained GIS professionals.

LITERATURE CITED

- American Society of Civil Engineers (ASCE), 1993, Survey of GIS applications in surface water hydrology *in* Engineering Hydrology, Chin Y. Kuo (ed.), American Society of Civil Engineers, New York, New York, p. 545-550.
- Anderson, J.R., Hardy, E.E., Roach, J.T., and Witmer, R.E., 1976, A land use and land cover classification system for use with remote sensor data: U.S. Geological Survey Professional Paper 964, 28 p.
- Battaglin, W.A., Kuhn, G., and Parker, R.S., 1993, Using a geographic information system to link digital spatial data and the precipitation-runoff modeling system in a hydro-climate study of the Gunnison River Basin, southwestern Colorado: Denver, Second International Conference on Integrating Geographic Information Systems, September 1993, 13 p.
- Benson, M.A., 1959, Channel-slope factor in flood frequency analysis: American Society of Civil Engineers, Separate no. 1994, 9 p.
- Benson, M.A., 1962a, Evolution of methods for evaluating the occurrence of floods: U.S. Geological Survey Water-Supply Paper 1580-A, 29 p.
- Benson, M.A., 1962b, Factors influencing the occurrence of floods in a humid region of diverse terrain: U.S. Geological Survey Water-Supply Paper 1580-B, 64 p.
- Benson, M.A., 1964, Factors affecting the occurrence of floods in the southwest: U.S. Geological Water Supply Paper 1580-D, 72 p.
- Berich, R.H., 1985, A micro-computer GIS for water resources planning: Buffalo, New York, Proceedings of the ASCE Conference on Computer Applications in Water Resources, p. 58-67.
- Berry, J., and Sailor, J., 1987, Use of geographical information systems for storm run-off prediction for small urban watersheds: Environmental Management, v. 11, no. 1, p. 21-27.
- Beven, K., and Moore, I., 1993, Introduction from terrain analysis and distributed modeling in hydrology: Chichester, New York, John Wiley and Sons, 249 p.
- Bhaskar, N.R., James, W.P., and Devulapalli, R.S., 1992, Hydrologic parameter estimation using geographic information system: American Society of Civil Engineers, Journal of Water Resources Planning and Management, v. 118, no. 5, p. 492-512.

- Brands, M.D., 1947, Flood runoff in the Willamette Valley, Oregon: U.S. Geological Survey Water Supply Paper 968-A, 59 p.
- Connors, K.F., Gardner, T.W., and Day, R.L., 1989, Deriving stream parameters for hydrologic modeling from digital elevation data —new directions for surface water modeling: Proceedings of the Baltimore Symposium, May 1989, IAHS publication no. 181, [unknown pagination].
- Cowden, R.W., 1991, Storm waste/wastewater management ---How an application can drive a GIS: New York, New York, American Society of Civil Engineers, Proceedings of the Water Resource Planning and Management and Urban Water Resource 18th Annual Conference, p. 913-917.
- Darling, M.E., and Hubbard, L.E., 1994, Application of a geographic information system for re-gridding a ground-water flow model of the Columbia Plateau regional aquifer system, Walla Walla River Basin, Oregon-Washington: U.S. Geological Survey Water Resources Investigations Report 89-4179, 23 p.
- Daly, C., and Neilson, R.P., 1992, A digital topographic approach to modeling the distribution of precipitation in mountainous terrain: Minneapolis, Minnesota, Interdisciplinary Approaches to Hydrology and Hydrogeology, American Institute of Hydrology, p. 437-454.
- Daly, C., Neilson, R.P., and Phillips, D.L., 1994, A statistical-topographic model for mapping climatological precipitation over mountainous terrain: Journal of Applied Meteorology, v. 33, no. 2, p. 140-158.
- DeVantier, B.A., and Feldman, A.D., 1993, Review of GIS applications in hydrologic modeling, Journal of Water Resources Planning and Management, v. 119, no. 2, p. 246-261.
- Dunne, T., and Leopold, L.B., 1978, Water in environmental planning: San Francisco, California, W. H. Freeman and Company, 818 p.
- Environmental Systems Research Institute (ESRI), 1990, Understanding GIS — The ARC/INFO method: Redlands, California, Environmental Systems Research Institute User's Manual, 10 chapters + appendices.
- Environmental Systems Research Institute (ESRI), 1991, Surface modeling with TIN —surface modeling and display: Redlands, California, Environmental Systems Research Institute User's Manual, 237 p.
- Environmental Systems Research Institute (ESRI), 1992, Map projections and coordinate management: Redlands, California, Environmental Systems Research Institute User's Manual, 86 p. +appendices.

- Fegeas, R.G., Claire, R.W., Guptill, S.C., Anderson, K.E., and Hallam, C.A., 1983, Land use and land cover digital data —U.S.Geological Survey Digital Cartographic Data Standards: U.S. Geological Survey Circular 895-E, 21 p.
- Fountain, A.G. and Tangborn, W., 1985, Overview of contemporary techniques in Techniques for prediction of runoff from glacierized areas: IAHS Publication no. 149, p. 27-41
- Gerig, A.J., 1985, Soil survey of Clackamas County area. Oregon: U.S. Department of Agriculture, Soil Conservation Service, 293 p.
- Goulter, I., and Forrest, D., 1987, Use of GIS in river basin management: Water Science Technology, v. 19, no. 9, p. 81-86.
- Green, G.L., 1983, Soil survey of Multnomah County, Oregon: U.S. Department of Agriculture, Soil Conservation Service, 225 p.
- Greis, N.P. and Wood, E.F., 1981, Regional flood frequency estimation and network design: Water Resource Research, v. 17, p. 1167-77.
- Gupta, A. and Fox, H., 1974, Effects of high-magnitude floods on channel form — A case study in Maryland piedmont: Water Resources Research, v. 10, p. 499-509.
- Harr, R.D., 1976, Forest practises and streamflow in western Oregon: U.S. Department of Agriculture, Forest Service General Technical Report PNW-49, Pacific Northwest Forest and Range Experiment Station, Portland, Oregon, 18 p.
- Harr, R. D., 1980, Hydrology of small forest streams in Western Oregon: U.S. Department of Agriculture Forest Service General Technical Report PNW-55, Pacific Northwest Forest and Range Experiment Station, Portland, Oregon, 15 p.
- Harr, R.D., Harper, W.C., Krygier, J.T., and Hsieh, F.S., 1975, Changes in storm hydrographs after road building and clear-cutting in the Oregon Coast Range: Water Resources Research, v. 11, no. 3, p. 436-444.
- Harris, D.D., and Hubbard, L.E., 1983, Magnitude and frequency of floods in eastern Oregon: U.S. Geological Survey Water Resources Investigations Report 82-4078, 39 p.
- Harris, D.D., Hubbard, L.L. and Hubbard, L.E., 1979, Magnitude and frequency of floods in western Oregon: U.S. Geological Survey Open-File Report 79-553, 35 p.

- Hay, L.E., Knapp, L., and Bromberg, J., 1993, Integrating geographic information systems, scientific visualization systems, statistics, and an orographic precipitation model for a hydro-climatic study of the Gunnison River Basin: Second International Conference on Integrating GIS & Environmental Modeling, Denver, Colorado, September 1993, 11 p.
- Hill, J.M., Singh, V.P., and Aminian, H., 1987, A computerized data base for flood prediction modeling: *Water Resource Bulletin*, v. 23, no. 1, p. 21-27.
- Hirschboeck, K.K., 1988, Flood hydroclimatology *in* Flood Geomorphology, eds. V. Baker, R. Kochel, and P. Patton, New York: Wiley, p. 27-49.
- Hobert, S.B., 1989, Development of a geographic information system based hydrologic model for stormwater management and landuse planning: Tuscon, Arizona, University of Arizona, M.S. thesis, [unknown pagination].
- Hubbard, L.E., Herrett, T.A., Kraus, R.L., Ruppert, G.P., and Courts, M.L., 1994, Water resources data, Oregon, water year 1993: U.S. Geological Survey Water-Data Report OR-93-1, 498 p.
- Hulsing, H. and Kallio, N.A., 1964, Pacific slope basins in Oregon and lower Columbia river basin, Part 14 of magnitude and frequency of floods in the United States: US Geological Survey Water Supply Paper 1689, 320 p.
- Jarvis, C.S. and others, 1936, Floods in the United States - magnitude and frequency: U.S. Geological Survey Water Supply Paper 771, 498 p.
- Jennings, M.E., Thomas, W.O., and Riggs, H.C., 1994, Nationwide summary of U.S. Geological Survey regional regression equations for estimating magnitude and frequency of floods for ungaged sites, 1993 —Reston, Virginia: U.S. Geological Survey Water Resources Investigations Report 94-4002, 196 p. + 1 computer disc
- Jenson, S.K., 1991, Applications of hydrologic information automatically extracted from digital elevation models: *Hydrologic Processes*, v. 5, no. 1, p. 31-44.
- Jenson, S.K., and Dominique, J.O., 1988, Extracting topographic structure from digital elevation data for geographic information system analysis: *Photogrammetric Engineering and Remote Sensing*, v. 22, no. 1, p. 15-24.
- Jeton, A.E., and Smith, J.L., 1993, Development of watershed models for two Sierra Nevada basins using a geographic information system: American Water Resources Association, *Geographic Information Systems and Water Resources*, A p. 251-258.

- Jones, N.L., Wright, S.G., and Maidment, D.R., 1990, Watershed delineation with triangle-based terrain models: American Society of Civil Engineers, Journal of Hydrologic Engineering, v. 116, no.10, p. 1232-1251.
- Kates, R.W., 1962, Hazard and choice perception in flood plain managment: University of Chicago, Department of Geography, Research Paper no. 78.
- Kirby, W., 1981, Annual flood frequency analysis using U.S. Water Resources Council Guidelines (Program J407), —Chapter I, Section C in Instructions for peak flow of WATSTORE user's guide: U.S. Geological Survey Open-File Report 79-1336-I, v. 4, p. C1- C57.
- Laenen, A., 1983, Storm runoff as related to urbanization based on data collected in Salem and Portland, and generalized for the Willamette Valley, Oregon: U.S. Geological Survey Water Resources Investigations Report 83-4143, 88 p.
- Laenen, A., 1995, Senior Hydrologist, U.S. Geological Survey, Water Resource Division Portland, Oregon, Personal communication.
- Laenen, A. and Solin, G.L., 1978, Rainfall-runoff data for selected basins, Portland, Oregon and Vancouver, Washington 1973-1977: U.S. Geological Survey Open File Report 78-291, 48 p.
- Langbein, W. B., 1947, Topographic characteristics of drainage basins: U. S. Geological Survey Water Supply Paper 968-C, p. 125-157.
- Leipnik, M.R., Kemp, K.K. and Loaiciga, H.A., 1993, Implementation of GIS for water resources planning and management: Journal of Water Resource Planning and Management, v. 119, no. 2, p. 184-205.
- Leipnik, M.R., and Loaiciga, H.A., 1991, Use of GIS in the study of water resources: Proceedings of the American Society of Agricultural Engineers, Water Resource Planning and Management and Urban Water Resource 18th Annual Conference and Symposium, p. 887-892.
- Lorenz, D.L., 1990, A method for describing stream-drainage-system topology in a geographic information system in B.H. Balthrop and E.G. Baker, compilers U.S. Geological Survey National Computer Technical Meeting Program and Abstracts, May 7-11, 1990: U.S. Geological Survey Open File Report 90-161, p. 22.
- Lystrom, D.J., 1970, Evaluation of the streamflow data program in Oregon: U.S. Geological Survey Unnumbered Open-File Report, 28 p. [available from U.S. Geological Survey, Oregon District, Portland, Oregon].

- Mallows, C.L., 1973, Some comments on Cp: *Technometrics*, v. 15, p. 661-675.
- Marks, D., Dozier, J., and Frew, J., 1984, Automated basin delineation from digital elevation data: *GEO-Processing*, v. 2, p. 299-311.
- Meyer, S.P. Salem, T.H. and Labadie, J.W., 1993, Geographic information systems in urban storm water management: *Journal of Water Resource Planning and Management*, v. 119, no. 2, p. 206-228.
- Moore, I.D., Grayson, R.B., and Ladson, A.R., 1993, Digital terrain modeling —A review of hydrological, geomorphological and biological applications *in* Beven, K., and Moore, I., 1993, Introduction from terrain analysis and distributed modeling in hydrology: Chichester, New York, John Wiley and Sons, p. 7-34.
- Muzik, I. and Pomeroy, S.J., 1990, A geographic information system for prediction of design flood hydrographs: *Canadian Journal of Civil Engineering*, v. 17, no. 6, p. 965-973.
- O'Callaghan, J. F., and Mark, D.M., 1984, The extraction of drainage data from digital elevation data: *Computer Vision, Graphics, and Image Process*, v. 28, no. 3, p. 323-344.
- Ott, L., 1988, An introduction to statistical methods and data analysis: Boston, Massachusetts, PWS-Kent Publishing Company, (3rd ed.), 835 p.
- Patterson, J.L., 1971, Floods in Arkansas, magnitude and frequency characteristics through 1968 —Arkansas Geologic Commission: U.S. Geological Survey Water Resources Circular no. 11, 199 p.
- Platt, R.H., 1986, Floods and Man —A geographer's agenda, Themes from the work of Gilbert F. White: *Geography, resources and environment*, v. 2, eds. R.W. Kates and I. Burton, Chicago, University of Chicago Press.
- Pentland, R.L., and Cuthbert, D.L., 1971, Operational hydrology for ungaged streams by the grid square technique: *Water Resources Research*, v. 7, no. 2, p. 283-291.
- Riggs, H.C., 1968, Some statistical tools in hydrology: *Techniques of Water Resources Investigations of the U. S. Geological Survey*, Book 4, Chapter A1, 39 p.
- Rothacher, J., 1971, Regimes of streamflow and their modification by logging *in* J. Morris, ed., *Forest Land Uses and Stream Environment*: Corvallis, Oregon State University, p. 40-54.

- Rothacher, J., 1973, Does harvest in west slope Douglas-Fir increase peak flow in small forested streams?: U. S. Department of Agriculture, Forest Service Paper PNW-163, 13 p.
- Schmidt, A.R., Romak, R.D., 1991, Development and organization of a geographic information system data base and its application to investigation of rainfall/runoff model parameters in Illinois: U. S. Geological Survey Open-File Report 91-209, 52 p.
- Schoolmaster, F.A. and Marr, P.G., 1992, Geographic information systems as a tool in water use data management: American Water Resources Association, Water Resources Bulletin, v. 28, no. 2, p. 331-336.
- See, R.B., Naftz, D.L., and Qualls, C.L., 1992, GIS-assisted regression analysis to identify sources of selenium in streams: American Water Resources Association, Water Resources Bulletin, v. 28, no. 2, p. 315-330.
- Shamsi, U.M., 1993, A GIS application to hydrology: New York, Proceedings of the Symposium Engineering Hydrology, American Society of Civil Engineers, p. 371 - 374.
- Silfer, A.T., Kinn, G.J. and Hassett, J.M., 1987, A geographical information system utilizing the triangulated network as a basis for hydrologic modeling — AutoCarto 8: Falls Church, Virginia, Proceedings of the International Symposium on Computer Assisted Cartography, American Society of Photogrammetry, [unknown pagination].
- Snyder, J.P., 1982, Map projections used by the U.S. Geological Survey: Geological Survey Bulletin 1532, second edition, 313 p.
- State of Oregon Water Resources Board, 1965, Lower Willamette River Basin: Salem, Oregon, 148 p.
- Statware, Inc., 1992, Chapter 6 in Statit statistics reference manual, release 3.0: Corvallis, Oregon, Statware Inc., 48 p.
- Stuebe, M.M., and Johnston, D.M., 1990, Runoff volume estimation using GIS techniques, Water Resources Bulletin, v. 26, no. 4, [unknown pagination].
- Tachikawa, Y., Shiiba, M., and Takasao, T., 1993, Development of a basin geomorphic information system using a TIN-DEM data structure: American Water Resources Association, Geographic Information Systems and Water Resources, p. 163 - 172.
- Taylor, G.H., 1993, Normal annual precipitation for the period 1961-90: Oregon State University, Oregon Climate Service, April 1993 Version, 1 sheet.

- Thomas, B.E., Hjalmarson, H.W., and Waltemeyer, S.D., 1994, Methods for estimating magnitude and frequency of floods in the southwestern United States: U.S. Geological Survey Open-File Report 93-419, 211 p.
- Thomas, D.M., and Benson, M.A., 1970, Generalization of streamflow characteristics from drainage-basin characteristics: U.S. Geological Survey Water-Supply Paper 1975, 55 p.
- U.S. Army Corps of Engineers, 1965, Flood of December 1964 in the Willamette River Basin, Oregon: Portland District, [pagination unknown].
- U.S. Fish and Wildlife Service, 1995, National wetland inventory digital data layers: Portland, Oregon, U.S. Fish and Wildlife Service, map scale 1:24,000.
- U. S. Water Resources Council, 1981, Guidelines for determining flood-flow frequency, Washington D.C., Water Resource Bulletin 17B, 28 p.
- Waylen, P.R., 1985, Stochastic flood analysis in a region of mixed generating processes: Transactions of the Institute of British Geographers, v. 10, p. 95-108.
- Waylen, P.R. and Woo, M.K., 1982, Prediction of annual floods generated by mixed processes: Water Resources Research, v. 18, p. 1283-1286.
- Weghorst, P.A., Cunningham, J. and Mortazavi, B., 1991, Geographic information systems in water resource management: Proceedings of the 1991 National Conference of Hydraulic Engineering, p. 888-893.
- Wellman, R.E., Gordon, J.M., and Moffatt, R.L, 1993, Statistical summaries of streamflow data in Oregon --Volume 2, Annual low and high flow, and instantaneous peak flow: U. S. Geological Survey Open-File Report 93-63, 406 p.
- White, G.F., 1945, Human adjustment to floods: University of Chicago, Department of Geography, Research Paper no. 29.
- White, G.F., 1964, Choice adjustment of floods: University of Chicago, Department of Geography, Research Paper no. 93.
- Whyte, A.V.T., 1986, From hazard perception to human ecology: Geography, resources and environment, eds. R.W. Kates and I. Burton, Chicago, University of Chicago Press, p. 240-271.
- Wolman, M.G. and Miller, J.P., 1960, Magnitude and frequency of forces in geomorphic processes: Journal of Geology, v. 68, p. 54-74.

- Woo, M.K. and Waylen, P.R., 1984, Areal prediction of annual floods generated by two distinct processes: *Hydrologic Sciences Journal*, v. 29, p. 75-88.
- Woo, M.K. and Waylen, P.R., 1986, Probability studies of floods, *Applied Geography*, v. 6, 185-195.
- Woodbury, M.S., and Jawed, K., 1993, Incorporating GIS technology in standard flood hydrology applications: New York, New York, American Society of Civil Engineers, *Proceedings of the Symposium in Engineering Hydrology*, p. 365 - 370.
- Wright, W.R., and Buehler, K.T., 1989, GIS - A proving ground for water resource management research: West Lafayette, Indiana, Purdue University, [unknown pagination].
- Zelt, R.B., 1991, GIS technology used to manage and analyze hydrologic information: *GIS WORLD*, v. 4, no. 5, p. 70.

APPENDIX A.

**PEAK DISCHARGE IN CUBIC FEET PER SECOND FOR SELECTED FLOOD
FREQUENCIES AT GAGED LOCATIONS.**

Station Number	0.50 (2-yr)	0.20 (5-yr)	0.10 (10-yr)	0.04 (25-yr)	0.02 (50-yr)	0.01 (100-yr)
14131000	76	122	158	212	259	310
14131400	304	376	421	475	514	916
14134000	283	451	583	777	940	1,121
14134500	1,361	2,093	2,642	3,406	4,027	4,691
14135000	5,359	7,523	9,108	11,289	13,046	14,921
14135500	4,780	7,264	9,185	11,941	14,244	16,769
14137000	14,500	22,112	27,403	34,293	39,541	44,872
14138800	1,096	1,471	1,710	2,002	2,214	2,421
14138850	5,790	7,000	7,737	8,615	9,239	9,841
14138870	569	786	939	1,144	1,306	1,474
14139700	996	1,377	1,614	1,896	2,095	2,286
14139800	1,751	2,340	2,715	3,175	3,508	3,834
14141500	2,157	3,043	3,648	4,432	5,030	5,639
14208000	3,008	4,749	6,031	7,782	9,176	10,643
14208500	504	713	853	1,033	1,168	1,303
14209000	1,649	2,433	3,017	3,830	4,491	5,201
14209500	17,141	25,148	30,382	36,858	41,571	46,184
14210000	24,636	37,676	46,564	57,921	66,416	74,911
14211000	38,827	55,008	67,300	84,736	99,178	114,928

APPENDIX B.

CONVERSION TABLE

Multiply	By	To Obtain
inch (in.)	25.40	millimeter (mm)
foot (ft)	0.3048	meter (m)
square mile (mi ²)	2.590	square kilometer (km ²)
cubic feet per second (ft ³)	0.02832	cubic meter per second (m ³ /s)
cubic feet per second per square mile [(ft ³ /s)/mi ²]	0.01093	cubic meter per second per square kilometer [(m ³ /s)/km ²]

Air temperatures are given in degrees Fahrenheit (°F), which can be converted to degrees Celsius (°C) by the following equation:

$$^{\circ}\text{C} = 5/9(^{\circ}\text{F}) - 32$$

APPENDIX C.

DRAINAGE BASIN CHARACTERISTIC VALUES USED IN THE REGRESSION ANALYSIS AS EXPLANATORY VARIABLES

Station Number	Drainage Area (mi ²)	Average Annual Precipitation (inches)	Water Storage (1:100,000) (decimal percent + .01)	Water Storage (1:24,000) (decimal percent + .01)	Forest Cover (decimal percent)	Area greater than 5000 feet (1-decimal percent)
14131000	3.78	98.9	0.01	0.0129	0.819	0.5967
14131400	14.68	98.3	0.01	0.0126	0.8444999	0.7219
14134000	8.00	79.9	0.01	0.0691	0.877	0.7639
14134500	52.53	65.6	0.017	0.0418	0.9753	0.9641
14135000	98.48	76.3	0.0138	0.0289	0.9861	0.9808
14135500	105.93	76.8	0.0136	0.0277	0.9783999	0.9822
14137000	259.26	85.9	0.0121	0.0238	0.9523	0.9524
14138800	8.17	92.1	0.012	0.0247	0.9969	1
14138850	48.33	107.6	0.0251	0.0279	0.9829	1
14138870	5.34	104.9	0.01	0.0229	1	1
14139700	7.90	101.4	0.01	0.0192	1	1
14139800	15.53	98.0	0.01	0.0153	1	1
14141500	23.84	85.9	0.0159	0.0110	0.9523	1
14208000	136.10	78.7	0.0111	0.0190	0.9985	0.9429
14208500	52.94	49.6	0.0264	0.0873	0.9437	0.9937
14209000	125.88	57.0	0.018	0.0474	0.9741	0.9973
14209500	488.70	72.8	0.0129	0.0256	0.9912	0.9809
14210000	679.87	74.5	0.0131	0.0239	0.9837	0.9861
14211000	933.94	71.3	0.0123	0.0236	0.8931	0.9989

APPENDIX C. (CONTINUED)

DRAINAGE BASIN CHARACTERISTIC VALUES USED IN THE REGRESSION ANALYSIS AS EXPLANATORY VARIABLES

Station Number	Main Channel Slope (ft / mi)	Main Channel Length (mi)	Total Upstream River Length (mi)	Area of Glaciers (decimal percent)	Mean Elevation (ft)	Lag Time (L / square root of S)
14131000	640	5.38	11.421	0.0907	4971.37	0.2126632
14131400	510	8.08	33.217	0.0503	4513.53	0.3577884
14134000	453	6.09	12.418	0.0471	4565.96	0.2861332
14134500	216	17.25	52.222	0.0072	3887.70	1.173714
14135000	182	27.10	111.936	0.0038	3432.88	2.008786
14135500	161	32.04	120.381	0.0036	3403.02	2.525106
14137000	137	40.37	357.675	0.0109	3319.28	3.449042
14138800	408	3.43	11.335	0	3348.42	0.1698103
14138850	232	12.90	62.888	0	3046.67	0.8469265
14138870	466	4.58	5.783	0	3055.38	0.2121644
14139700	291	5.48	8.855	0	2910.47	0.3212433
14139800	251	9.54	20.204	0	2695.99	0.6021594
14141500	206	13.76	23.150	0	2410.05	0.9587047
14208000	131	21.54	176.500	0	3929.05	1.881959
14208500	103	9.54	44.656	0	3751.07	0.9400041
14209000	104	19.15	115.990	0	3721.81	1.877812
14209500	97	38.64	575.482	0	3551.95	3.923297
14210000	74	62.61	792.267	0	3340.67	7.278265
14211000	60	80.06	1097.110	0	2755.16	10.3357

APPENDIX D.

GLOSSARY OF SELECTED TERMINOLOGY

- Adjusted R^2 .** A correlation coefficient that allows comparisons among models with differing numbers of explanatory variables, adjusts for the degrees of freedom in the model, and penalizes a model that includes excess slope parameters.
- Akaike' information criterion (AIC).** A statistical test similar to Mallow's C_p that includes a measure of model error and a penalty for excess variables. Models with small AIC are preferred.
- ADAPS.** Automated Data Processing System (ADAPS) consists of a collection of computer programs and files designed to provide, standardized water data processing procedures. ADAPS information is used to compute streamflow, reservoir, or other types of hydrologic records and can be displayed in table or graphical format.
- arc attribute table.** A table containing attributes for a line digital data file.
- ARC macro language (AML).** A Geographic Information System programming language specific to the ARC/INFO GIS software that provides full computer programming capabilities for spatial data analyses and turn key applications.
- arc.** A digital representation of a line segment using a string of x,y coordinate pairs (vertices). For example, a stream channel can be represented by one or more arc segments that are made up a series of geographic x,y coordinates.
- ARC/INFO¹.** A Geographical Information System (GIS) that is commercially available from Environmental Systems Research Institute (ESRI) in Redlands, California. ARC/INFO has two major subsystems, ARC and INFO. ARC is comprised of utilities used to create, manage, analyze and graphically display geographic digital data. INFO is a relational data-base management system (DBMS) which is utilized by the ARC subsystem for the storage of registration tics, map boundaries, and attribute information.

¹The use of brand names in this report is for identification purposes only and does not constitute an endorsement by the author.

APPENDIX D. (CONTINUED)

ASCII flat file. ASCII is the acronym used for the American Standard Code for Information Interchange. An ASCII flat file is a system file made up of a set of codes representing alphanumeric information (ie., a byte with a value of 77 represents a capital M). Text files, such as those created with a text editor of a computer system, are often referred to as ASCII files.

attribute. Information or a characteristic, (such as name, type, length) which describes a geographic feature (such as a stream channel) in a digital data layer. Attribute information is typically stored in tabular format, and is linked to the geographic feature by a user defined identifier.

clip. The spatial extraction of features contained in one data layer that fall entirely within a boundary defined by features in another data layer (often referred to as a clip coverage). A clip coverage works much like a *cookie cutter* by cutting out the features of a data coverage using the outer boundary of the clip coverage.

coverage. "Digital data layer" and "coverage" are used interchangeably in this report. A coverage (or data layer) is a digital version of a map consisting of geographically referenced features usually representing a single map theme such as a roads, streams, soils or land use.

data layer. See *coverage*.

digital elevation model (DEM). A elevation data base representing topographic relief by a set of regularly spaced x,y,z coordinates where z represents surface elevation. Digital elevation models provide a digital representation of a continuous variable over a two-dimensional surface by a regular array of z values referenced to a common datum.

digitize. To encode map features as x,y coordinates and store in a digital format using a digitizer.

digitizer. A device that consists of a digitizing table or tablet and a cursor with cross-hairs and keys used to record the location of map features as x,y coordinates.

APPENDIX D. (CONTINUED)

drainage basin. (also called "watershed" or "catchment") The area of land that drains water, sediment, and dissolved materials to a common outlet at some point along a stream channel.

exceedance probability. Probability that a random event will exceed a specific magnitude in a given time period. For example, a flood with a 0.01 exceedance probability is a flood that has one chance in a hundred of being exceed in any one year. This is a 100-year flood under the "recurrence interval" terminology (see recurrence interval). In this report, the term "recurrence interval" is used in preference to the term "exceedance probability". Both terms, however, are used in most of the tables, graphs, and illustrative problems.

"F" statistic. A test that defines the significance of the independent variables. The larger the "F" value, the more significance it has in the equation.

file transfer protocols (FTP's). FTP's are used to transport files from remote locations.

item. A field of information in an attribute table, displayed as a column.

hydrologic unit. Drainage basins located throughout the United States, determined by the topography and used to provide standard framework for water-resource planning.

geographical information system (GIS). An organized collection of computer hardware, software, geographic data, and personnel designed to efficiently capture, store, update, manipulate, analyze, and display all forms of geographically referenced information.

geo-reference. To establish the relationship between coordinates on a planar map and known real-world coordinates

lag time. The time from beginning or center of mass of rainfall to peak or center of mass of runoff. Lag time is a function of the main channel length divided by the square root of the main channel slope.

lattice. Lattices are the surface interpretation of a grid, that uses a rectangular array of points spaced at a constant sampling interval in the x and y directions relative to a common origin. A lattice is stored as a grid, but differs from a grid

APPENDIX D. (CONTINUED)

in that it represents the value of the surface only at the mesh points of the lattice rather than the value of the cell area surrounding each mesh point. In ARC/INFO terminology, a lattice is a grid.

line topology. A set of ordered coordinates that represent a linear feature such as roads, streams and contours.

macro. A text file containing a series of commands that can be executed as one command. Macros are designed to perform multiple operations such as spatial overlay processes using several data layers, or multiple iterative computations of area and length values. The ARC macro language (AML) is used to create macros for ARC/INFO.

Mallows' C_p . Proposed by C.P. Mallows, this statistics is used as an aid in choosing a final regression model. It is a measure of total squared error for a subset model containing p independent variables. The total squared error is a measure of the error variance plus the bias introduced by not including significant variables in a model. The equation for Mallows' C_p is in the form of:

$$C_p = (SSE(p)/MSE) - (N-2p) + 1$$

where MSE = mean square error, $SSE(p)$ = sum of squares error, N = total sample size.

map projection. A systematic conversion of locations of the Earth's surface from spherical to planar coordinates using a mathematical model.

overlay. Process which merges two overlapping coverages and their attributes to form a third coverage.

Pearson's r . A measure of linear correlation. It is also called the linear correlation coefficient because r measures the linear association between two variables.

point or polygon attribute table (PAT). A point or polygon attribute table, contains attribute data for a point or polygon coverage. These data base management files are used to store coverage attribute information. ARC/INFO automatically writes the following information to attribute tables of Points or Polygons: AREA, PERIMETER, COVER# and COVER-ID (also known as the feature User-ID).

APPENDIX D. (CONTINUED)

point topology. Feature in a coverage represented by a single x,y coordinate.

polygon topology. An areal feature in a coverage defined by the lines which bound it.

projection. A systematic conversion of locations on the earth's surface from spherical to planar coordinates.

R-square. The coefficient of determination, or a measure of the variation explained by the regression equation. R-square \times 100 yields the percent of variation explained. If R-square = 1, then 100 percent of the variation is explained by the equation; if R-square = 0.75, then 75 percent of the variation is explained by the equation. R-square is a "best-fit" test for the population scatter about the curve.

relate. A temporary link between records in two files based on an item common to both.

root mean square error (RMS error). Root Mean Square error is primarily used in GIS for tic registration and represents the deviation or amount of error between the tic locations in the original and the new coordinate locations. The mathematical expression for computing an RMS error is provided in ARC/INFO 6.0 user's manual "Map Projections & Coordinate Management" (ESRI, 1992). A perfect transformation would produce an RMS error of 0.000, however this is difficult to obtain and most applications except an RMS error of 0.004 or below.

standard error of estimate (SEE). A statistical measure of accuracy based on the population scatter about the curve only. It is the square root of the variance and is graphically defined as having two-thirds of the data points falling within its limits.

APPENDIX E. ARC MACRO LANGUAGE PROGRAMS

```

/*****
/***** AML PROGRAM ARCSUM.AML *****/
/*****
/* Version:                1.0
/* Language:               Arc Macro Language
/* Arc Version/Platform:   6.0 / DGUX
/* Subsystem:              Arc
/*
/*.....
/* Purpose: Computes total upstream channel length for all streams located above a selected
/*          stream site. Transfers attributes from a line coverage (hydrography) to a point
/*          coverage (sampling sites).
/*.....
/* Arguments:
/* Variable name      Type      Definition
/* -----
/* itemname          character   Name of item in AAT to be transferred.
/* ptcov              character   Name of point coverage.
/* hydrobasincov      character   Name of hydrography and basin delineation coverage.
/*.....
/* Programs or menus called: none
/*.....
/* History:
/* Author/Site        Date        Version      Event
/* -----
/* Harrison/USGS-OR 6-93          1.0          Original coding
/* Brownell/USGS-OR 6-94          1.0          Revisions
/*.....
/* Disclaimer:
/* This program has been used by the U.S. Geological Survey; no warranty, either expressed
/* or implied, is made by the USGS as to the accuracy or functioning of the program.
/* No responsibility is assumed by the USGS in connection with this program's distribution or
/* use.
/*-----
&args itemname ptcov hydrobasincov
&severity &error &routine get_out
&select [locase [show program]]
&when arc
&do
&type
&type
&type //////////////////////////////////////
&type
&type Disclaimer:
&type

```

```

&type This program has been used by the U.S. Geological Survey;
&type no warranty, either expressed or implied, is made by the USGS as to the
&type accuracy or functioning of the program. No responsibility is assumed
&type by the USGS in connection with this program's distribution or use.
&type
&type //////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////
&type
&type ARCSUM.AML : Version 1.0 : [date -full]
&type
&s tube [extract 1 [show &term]]
&if [null %tube%] &then
&do
&term 9999 &mouse
display 9999 position ul screen
&end
&end
&otherwise
&do
&type This program is run from ARC prompt ...
&stop
&end
&end
/*-----
/* Check for arguments
/*-----
&if [null %itemname%] &then &call usage
&if [null %ptcov%] &then &call usage
&if [null %hydrobasincov%] &then &call usage
/*-----
/* Remove previous arcsum.list file
/*-----
&if [exists %itemname%sum.list -file] &then ~
&s delstat [delete %itemname%sum.list -file]
/*-----
/* Remove previous tracelist file
/*-----
&if [exists tracelist -file] &then &s delstat [delete tracelist -file]
/*-----
/* Check for site coverage
/*-----
&if ^ [exist %ptcov% -cover] &then
&do
&type %ptcov% coverage NOT found
&stop
&end
/*-----
/* Check for hydro/basin coverage
/*-----
&if ^ [exist %hydrobasincov% -cover] &then
&do
&type %hydrobasincov% coverage NOT found

```

```

&stop
&end
/*-----
/* Check for item to be summed in sites coverage
/*-----
&s dfn = 0
&if ^ [iteminfo %ptcov% -point %itemname% -exists] &then
&do
&if [iteminfo %hydrobasincov% -line %itemname% -exists] &then
/*-----
/* Item not found in site coverage but in hydro/basin coverage
/*-----
&do
&set dfn [iteminfo %hydrobasincov% -line %itemname% -definition]
additem %ptcov%.pat %ptcov%.pat %itemname% %dfn%
&end
&else
/*-----
/* Item not found in hydro/basin or site coverage
/*-----
&do
&type PROGRAM FAILURE
&type %itemname% not found in %hydrobasincov%
&stop
&end
&end
/*-----
/* Set definition variable if not set above
/*-----
&if x%dfn% = x0 &then ~
&set dfn [iteminfo %ptcov% -point %itemname% -definition]
/*-----
/* Add flag item to hydro/basin coverage
/*-----
&if ^ [iteminfo %hydrobasincov% -line arcflag -exists] &then ~
additem %hydrobasincov%.aat %hydrobasincov%.aat arcflag 1 1 I
/*-----
/* Count readfiles and place into tracelist file for loop
/*-----
&s count [filelist *.trace tracelist -file]
&if %count% le 0 &then
&do
&type PROGRAM FAILURE
&type filelist has returned %count%
&return
&end
/*-----
/* Begin loop to read readflies from tracelist
/*-----
&s filunit [open tracelist openstatus -read]
&if %openstatus% ne 0 &then
&do

```

```

&type PROGRAM FAILURE
&type tracelist not open
&type open has returned %openstatus%
&return
&end
&do &until %readstatus% = 102
&s readfile [read %filunit% readstatus]
&if %readstatus% eq 0 &then
&do
/*-----
/* Identify site being processed
/*-----
&s siteno [subst %readfile% .trace]
&type
&type
&type =====
&type
&type PROCESSING SITE %siteno%
&type
&type =====
&type
&type
/*-----
/* Flag reaches upstream from site
/*-----
arcplot
display 1040
testplot
mape %hydrobasincov%
calculate %hydrobasincov% line arcflag = 0
readselect %readfile%
calculate %hydrobasincov% line arcflag = 1
quit /*ap
/*-----
/* Sum item for reaches upstream from site
/*-----
&data arc info
ARC
SEL [translate %hydrobasincov%].AAT
RES $RECNO = 1
FORMAT $NUM1, %dfn%
FORMAT $NUM5,8,I
CALC $NUM1 = 0
ASEL
RES ARCFLAG = 1
CALC $NUM5 = $NOSEL
CALC $NUM1 = $NUM1 + [translate %itemname%]
REM
REM Transfer sum value to item in site coverage
REM
SEL [translate %ptcov%].PAT
RES STATION-ID = %siteno%

```

```

CALC [translate %itemname%] = $NUM1
OUTPUT ../[translate %itemname%]SUM.LIST APP
CALC $COMMA-SWITCH = -1
PRINT 1T,'THERE ARE ', $NUM5,' ARCS ABOVE SITE ',STATION-ID
PRINT 1T,'SUM ',[quote [translate %itemname%]],2X,'=',2X,[translate %itemname%]
PRINT 80X
Q STOP
&end /* &data
&end /* &do
&end /* &do &until
&s closeststatus [close %filunit%]
&type PROGRAM COMPLETE
&return
/*
&routine get_out
&type PROGRAM HAS FAILED
&type
&s closeststatus [close -all]
&stop
/*
&routine usage
&type USAGE: ARCSUM <aat_item_name> <sites_coverage> <hydro_basin_cov>
&stop
*****
*****
*****

```

```

/*****
/***** AML PROGRAM MCLENGTH.AML *****/
/*****
/* Version:          1.0
/* Language:         Arc Macro Language
/* Arc Version/Platform:  6.0 / DGUX
/* Subsystem:        Arc
/*
/
*.....
/* Purpose: To calculate the length of the longest arc (stream reach) upstream from a selected
/*          node (gaging site).
/*.....
/* Variable name      Type      Definition
/*-----
/* itemname           character   Name of item in AAT to be transferred
/* ptcov              character   Name of point coverage
/* hydrobasincov      character   Name of hydrography and basin delineation coverage
/*.....
/* Programs or menus called: none
/*.....
/* Author/Site        Date   Version   Event
/* -----
/* Harrison/USGS-OR   6-93    1.0       Original coding
/* Harrison/USGS-OR   5-94    1.0       Update
/*.....
/* Disclaimer:
/* Although this program has been used by the U.S. Geological Survey, no warranty, expressed
/* or implied, is made by the USGS as to the accuracy or functioning of the program.
/* No responsibility is assumed by the USGS in connection with this program's distribution or
/* use.
/*-----
/*
&args itemname ptcov hydrobasincov
&severity &error &routine get_out
&select [locase [show program]]
&when arc
&do
&type
&type
&type //////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////
&type
&type Disclaimer:
&type
&type This program has been used by the U.S. Geological Survey;
&type no warranty, expressed or implied, is made by the USGS as to the
&type accuracy or functioning of the program. No responsibility is assumed
&type by the USGS in connection with this program's distribution or use.
&type
&type //////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////
&type
&type MCLENGTH.AML : Version 1.0 : [date -full]

```



```

&type
&s tube [extract 1 [show &term]]
&if [null %tube%] &then
  &do
    &term 9999 &mouse
    display 9999 position ul screen
  &end
&end
&otherwise
  &do
    &type This program is run from ARC prompt ...
    &stop
  &end
&end
/*-----
/* Check for args
/*-----
&if [null %itemname%] &then &call usage
&if [null %ptcov%] &then &call usage
&if [null %hydrobasincov%] &then &call usage
/*
/*-----
/* Remove previous tracelist file
/*-----
/*
&if [exists tracelist -file] &then &s delstat [delete tracelist -file]
/*
/*-----
/* Check for site coverage
/*-----
&if ^ [exist %ptcov% -cover] &then
  &do
    &type %ptcov% coverage NOT found
    &stop
  &end
/*-----
/* Check for hydro/basin coverage
/*-----
&if ^ [exist %hydrobasincov% -cover] &then
  &do
    &type %hydrobasincov% coverage NOT found
    &stop
  &end
/*-----
/* Check for item in sites coverage
/*-----
&s dfn = 0
&if ^ [iteminfo %ptcov% -point %itemname% -exists] &then
  &do
    &if [iteminfo %hydrobasincov% -node %itemname% -exists] &then
      /*-----
      /* Item not found in site coverage but in hydro/basin coverage

```

```

/*-----
&do
  &set dfn [iteminfo %hydrobasincov% -node %itemname% -definition]
  additem %ptcov%.pat %ptcov%.pat %itemname% %dfn%
&end
&else
/*-----
/* Item not found in hydro/basin or site coverage
/*-----
&do
  &type PROGRAM FAILURE
  &type %itemname% not found in %hydrobasincov%
  &stop
&end
&end
/*-----
/* Set definition variable if not set above
/*-----
&if x%dfn% = x0 &then ~
&set dfn [iteminfo %ptcov% -point %itemname% -definition]
/*
/*-----
/* Count readselect files and place into tracelist file for loop
/*-----
&s count [filelist *.trace tracelist -file]
&if %count% le 0 &then
  &do
    &type PROGRAM FAILURE
    &type filelist has returned %count%
    &return
  &end
/*-----
/* Begin loop to read readselect files from tracelist
/*-----
&s filunit1 [open tracelist openstatus -read]
&if %openstatus% ne 0 &then
  &do
    &type PROGRAM FAILURE
    &type tracelist not open
    &type open has returned %openstatus%
    &return
  &end
&do &until %readstatus% = 102
  &s readfile [read %filunit1% readstatus]
  &if %readstatus% eq 0 &then
    &do
      /*-----
      /* Identify site being processed
      /*-----
      &s siteno [subst %readfile% .trace]
      &type
      &type

```

```

&type =====
&type
&type      PROCESSING SITE %siteno%
&type
&type =====
&type
&type
/*-----
/* initiate length and id variables
/*-----
&s sumlength = 0
&s maxlength = 0
&s arcid = 0
&s maxarcid = 0
/*
/*-----
/* read readselect files in AP (AE doesn't readselect)
/*-----
arcplot
mape %hydrobasincov%
/*-----
/* initialize arcflag
/*-----
calculate %hydrobasincov% arcs arcflag = 0
/*-----
/* flag all arcs above site with 2
/*-----
readselect %readfile%
calculate %hydrobasincov% arcs arcflag = 2
/*-----
/* flag headwater arcs with 1
/*-----
reselect %hydrobasincov% arcs upntr1 = 0
calculate %hydrobasincov% arcs arcflag = 1
q
/*-----
/* draw up hydrography in AE
/*-----
arcedit
editcov %hydrobasincov%
drawe arc
draw
ef arc
select arcflag = 2 or arcflag = 1
calc $symbol = 5
/*-----
/* set site coordinates
/*-----
editfeature node
select station-id = %siteno%
&s sitetnode = [show node [show select 1] id]
ef arc

```

```

sel bashyd cn 'H'
res tnode# = %sitetnode%
&s sitefnode = [show arc [show select 1] fnode#]
ef node
sel %hydrobasincov%-id = %sitefnode%
&s coordsite = [show node [show select 1] coordinate]
&type \\\
&type Site %siteno%'s coordinates are
&type [show node [show select 1] coordinate]
&type \\\
/*-----
/* select all headwater arcs and set counter to number selected
/*-----
editfeature arc
&s counter = 0
&label nextarc
select arcflag = 1
/*calc $symbol = 7
&s totalnum [show number select]
/*-----
/* increment counter
/*-----
&s counter = %counter% + 1
&type \\\
&type Working on headwater arc %counter% of %totalnum% ...
&type \\\
/*-----
/* select one headwater arc
/*-----
reselect %hydrobasincov%-id = [show arc [show select %counter%] id]
/*-----
/* capture arc id
/*-----
&s arcid = [show arc [show select 1] id]
/*-----
/* set coordinates of headwater arc selected
/*-----
&s coordhdwtrs = [show node [show arc [show select 1] fnode#] coordinate]
/*-----
/* select only hydrography arcs to prevent path on delineation
/*-----
select bashyd cn 'H'
/*-----
/* open watchfile to capture length of path
/*-----
&watch longest
/*-----
/* select path from headwaters to site
/*-----
coord keyboard
resel path
[unquote %coordhdwtrs%]

```

```

[unquote %coordsite%]
/*
/* close watch file
/*
&watch &off
/*-----
/* retrieve path length from watchfile
/*-----
&s filunit2 [open longest openstatus -read]
&if %openstatus% ne 0 &then ~
&type watchfile not opened successfully
/*-----
/* read first six lines
/*-----
&do linenum = 1 &to 6
  &type reading line number %linenum%
  &s dummy [read %filunit2% readstatus]
&end
&s string [read %filunit2% readstatus]
&s closestatus [close %filunit2%]
&s delerr [delete longest -file]
&type \\\
&type Here is line 7 of file
&type %string%
&s sumlength = [extract 10 [unquote %string%]]
&type sumlength is %sumlength%
&type \\\
&type working on %counter% out of %totalnum% headwater arcs
&type above site %siteno%
&type cumulative length = %sumlength%
&type current arc id = %arcid%
&type \\\
/*-----
/* compare new path length to previous path length
/*-----
&if %sumlength%x gt %maxlength%x &then &s maxarcid = %arcid%
&s maxlength = [max [unquote %maxlength%] [unquote %sumlength%]]
&type \\\
/*-----
/* if not last headwaters arc then go back to get another arc
/*-----
&if %counter% ne %totalnum% &then &goto nextarc
/*-----
/* move maxlength and maxarcid to point coverage
/*-----
&type Last headwaters reach has been measured.
&type Maximum channel length for %siteno% is %maxlength% meters
&type ID of arc farthest from site is %maxarcid%
&type \\\
&type Length and arc id are being written to site in point coverage ...
&type \\\
/*-----

```

```

/* write maximum channel length into site point coverage
/*-----
editcov %ptcov%
ef point
sel station-id = %siteno%
calc %itemname% = %maxlength%
calc mcarcid = %maxarcid%
calc %itemname%2 = %itemname% / 1609.2
save
q n
/*-----
/* get next site
/*-----
&end /*&do
&end /*&do until
&s closestatus [close %filunit1%]
&type PROGRAM COMPLETE
&return
/*
&routine get_out
&type PROGRAM HAS FAILED
&type
&s closestatus [close -all]
&stop
/*
&routine usage
&type Usage: MCLENGTH.AML <item_name> <point_cov> <hydro_basin_cov>
&stop
/*****
/*****
/*****

```

```

/*****
/***** AML PROGRAM CALCLU.AML *****/
/*****
/* Version:          1.0
/* Language:         Arc Macro Language
/* Arc Version/Platform:  6.0 / DGUX
/* Subsystem:        Arc
/
*.....
/* Purpose: To populate an INFO file with land use percentages and areas for entire drainage
/*          area located upstream from each sampling site.
/*.....
/* Before running:
/* 1. define INFO file SITES.LNDUSE (or purge records if rerunning).
/* 2. create ASCII file with list of sites delineated for which land use is to be calculated.
/*.....
/* Programs or menus called: None
/*.....
/* Author/Site      Date   Version  Event
/* -----
/* HARRISON/USGS-OR   8-93   1.0      Original coding
/* BROWNELL/USGS-OR   6-94   1.0      Update
/*
/*.....
/* Disclaimer:
/* This program has been used by the U.S. Geological Survey; no warranty, either expressed or
/* implied, is made by the USGS as to the accuracy or functioning of the program.
/* No responsibility is assumed by the USGS in connection with this program's distribution or
/* use.
/*-----
/*
&severity &error &routine get_out
&select [locase [show program]]
&when arc
&do
&type
&type
&type //////////////////////////////////////
&type
&type Disclaimer:
&type
&type This program has been used by the U.S. Geological Survey;
&type no warranty, either expressed or implied, is made by the USGS as to the
&type accuracy or functioning of the program. No responsibility is assumed
&type by the USGS in connection with this program's distribution or use.
&type
&type //////////////////////////////////////
&type
&type CALCLU.AML : Version 1.0 : [date -full]
&type
&s tube [extract 1 [show &term]]
&if [null %tube%] &then

```

```

    &do
        &term 9999 &mouse
        display 9999 position ul screen
    &end
&end
&otherwise
    &do
        &type This program is run from ARC prompt ...
        &stop
    &end
&end
/*
&if ^ [exists delinlist -file] &then &do
    &type There is no ASCII file (delinlist) containing basin names ...
    &stop
&end
&if ^ [exists sites.lnduse -info] &then &do
    &type There is no INFO file (sites.lnduse) to write into ...
    &stop
&end
&if [query 'Purge previous records in SITES.LNDUSE'] = .TRUE. &then &do
    &data arc info
ARC
SEL SITES.LNDUSE
ASEL
PURGE
Y
Q STOP
    &end
&end
&s fileu [open delinlist oerr -read]
&s rerr = 0
&do &until %rerr% = 102
    &s sitename [read %fileu% rerr]
    &if %rerr% ne 102 &then &do
        &type
        &type -----
        &type
        &type Calculating landuse percentages for: %sitename%
        &type
        &type -----
        &type
        &if ^ [exists %sitename% -cover] &then &goto nocov
        identity %sitename% landuse_net temp poly
        &describe temp
        &if %dsc$polygons% = 1 &then &goto nocov
        &data arc info
ARC
FO $NUM1,4,12,F,3
FO $NUM2,4,12,F,3
FO $NUM4,4,12,F,3
FO $NUM6,4,12,F,3

```



```

FO $NUM8,4,12,F,3
FO $NUM10,4,12,F,3
FO $NUM12,4,12,F,3
FO $NUM14,4,12,F,3
FO $NUM16,4,12,F,3
FO $NUM18,4,12,F,3
FO $NUM20,4,12,F,3
FO $NUM3,6,N,2
FO $NUM5,6,N,2
FO $NUM7,6,N,2
FO $NUM9,6,N,2
FO $NUM11,6,N,2
FO $NUM13,6,N,2
FO $NUM15,6,N,2
FO $NUM17,6,N,2
FO $NUM19,6,N,2
SEL TEMP.PAT
RES AREA GE 0
CALC $NUM1 = AREA + $NUM1
CALC $NUM20 = $NUM1 / ( 1609.2 ** 2 )
CALC $NUM1 = $NUM20
REM -----
RES LUC = 10
CALC $NUM2 = AREA + $NUM2
CALC $NUM20 = $NUM2 / ( 1609.2 ** 2 )
CALC $NUM2 = $NUM20
CALC $NUM3 = ( $NUM2 / $NUM1 ) * 100
ASEL
REM -----
RES LUC = 20
CALC $NUM4 = AREA + $NUM4
CALC $NUM20 = $NUM4 / ( 1609.2 ** 2 )
CALC $NUM4 = $NUM20
CALC $NUM5 = ( $NUM4 / $NUM1 ) * 100
ASEL
REM -----
RES LUC = 30
CALC $NUM6 = AREA + $NUM6
CALC $NUM20 = $NUM6 / ( 1609.2 ** 2 )
CALC $NUM6 = $NUM20
CALC $NUM7 = ( $NUM6 / $NUM1 ) * 100
ASEL
REM -----
RES LUC = 40
CALC $NUM8 = AREA + $NUM8
CALC $NUM20 = $NUM8 / ( 1609.2 ** 2 )
CALC $NUM8 = $NUM20
CALC $NUM9 = ( $NUM8 / $NUM1 ) * 100
ASEL
REM -----
RES LUC = 50
CALC $NUM10 = AREA + $NUM10

```



```

[unquote ' ']
[unquote ' ']
[unquote ' ']
[unquote ' ']
ASEL
RES NAME CN [quote [translate %sitename%]]
CALC DRAINAGE_AREA = $NUM1
CALC URBAN_AREA = $NUM2
CALC URBAN_PCT = $NUM3
CALC AGRIC_AREA = $NUM4
CALC AGRIC_PCT = $NUM5
CALC RANGE_AREA = $NUM6
CALC RANGE_PCT = $NUM7
CALC FOREST_AREA = $NUM8
CALC FOREST_PCT = $NUM9
CALC WATER_AREA = $NUM10
CALC WATER_PCT = $NUM11
CALC WETLND_AREA = $NUM12
CALC WETLND_PCT = $NUM13
CALC BARE_AREA = $NUM14
CALC BARE_PCT = $NUM15
CALC TUNDRA_AREA = $NUM16
CALC TUNDRA_PCT = $NUM17
CALC GLACIAL_AREA = $NUM18
CALC GLACIAL_PCT = $NUM19
Q STOP
    &end      /* for info data block
    &label nocov
    &if [exists temp -cov] &then kill temp
    &end      /* for do
&end      /* for do until
&s cerr [close %fileu%]
&type
&type ALL BASINS COMPLETED ...
&type
&return
/*
&routine get_out
&s cerr [close -all]
&if [exists temp -cov] &then kill temp
&type Program has failed ...
&stop
/*****
/*****
/*****

```

```

/*****
/***** AML PROGRAM CALCSTOR.AML *****/
/*****
/* Version:          1.0
/* Language:         Arc Macro Language
/* Arc Version/Platform: 6.0 / DGUX
/* Subsystem:        Arc
/*
/
*.....
/*
/* Purpose: To populate an INFO file with lake, pond, wetland, and reservoir area percentages
/*           for the entire drainage area located upstream from selected stream site.
/*
/*.....
/*
/* Arguments: None
/*
/* Before running:
/* 1. define INFO file SITES.LAKE (or purge records if rerunning).
/* 2. create ASCII file with list of sites delineated for which areas and percentages are to be
/*    calculated
/*
/*.....
/*
/* Programs or menus called: None
/*
/*.....
/*
/* History:
/*
/* Author/Site          Date   Version  Event
/* -----
/* HARRISON/USGS-OR    8-5-93  1.0    Original coding
/* BROWNELL/USGS-OR   6-15-94  1.0    Update
/*
/*.....
/*
/* Disclaimer:
/* This program has been used by the U.S. Geological Survey; no warranty, either expressed or
/* implied, is made by the USGS as to the accuracy or functioning of the program.
/* No responsibility is assumed by the USGS in connection with this program's distribution
/* or use.
/*
/*-----
/*
&severity &error &routine get_out
&select [locase [show program]]
&when arc
&do
&type
&type

```

```

&type //////////////////////////////////////
&type
&type Disclaimer:
&type
&type This program has been used by the U.S. Geological Survey;
&type no warranty, either expressed or implied, is made by the USGS as to the
&type accuracy or functioning of the program. No responsibility is assumed
&type by the USGS in connection with this program's distribution or use.
&type
&type //////////////////////////////////////
&type
&type CALCLAKE.AML : Version 1.0 : [date -full]
&type
&s tube [extract 1 [show &term]]
&if [null %tube%] &then
  &do
    &term 9999 &mouse
    display 9999 position ul screen
  &end
&end
&otherwise
&do
  &type This program is run from ARC prompt ...
&stop
&end
&end /* for select
/*
&if ^ [exists delinlist -file] &then &do
&type There is no ASCII file (delinlist) containing basin names ...
&stop
&end
&if ^ [exists sites.lake -info] &then &do
&type There is no INFO file (sites.lake) to write into ...
&stop
&end
&if [query 'Purge previous records in SITES.LAKE'] = .TRUE. &then &do
  &data arc info
ARC
SEL SITES.LAKE
ASEL
PURGE
Y
Q STOP
&end
&end
&s fileu [open delinlist oerr -read]
&s rerr = 0
&do &until %rerr% = 102
  &s sitename [read %fileu% rerr]
  &if %rerr% ne 102 &then &do
    &type
    &type -----

```

```

&type
&type Calculating lake area and percentages for: %sitename%
&type
&type -----
&type
&if ^ [exists %sitename% -cover] &then &goto nocov
identity %sitename% lakes temp poly
&describe temp
&if %dsc$polygons% = 1 &then &goto nocov
&data arc info
ARC
FO $NUM1,4,12,F,4
FO $NUM33,4,12,F,4
FO $NUM35,4,12,F,4
FO $NUM20,4,12,F,4
SEL TEMP.PAT
RES AREA GE 0
CALC $NUM1 = AREA + $NUM1
DISPLAY $NUM1
CALC $NUM20 = $NUM1 / ( 1609.2 ** 2 )
DISPLAY $NUM20
CALC $NUM1 = $NUM20
DISPLAY $NUM1
REM -----
RES MINOR1 = 421
CALC $NUM33 = AREA + $NUM33
DISPLAY $NUM33
CALC $NUM20 = $NUM33 / ( 1609.2 ** 2 )
DISPLAY $NUM20
CALC $NUM33 = $NUM20
DISPLAY $NUM33
CALC $NUM35 = ( $NUM33 / $NUM1 ) * 100
DISPLAY $NUM35
ASEL
REM -----
RES MINOR1 = 111
CALC $NUM37 = AREA + $NUM37
DISPLAY $NUM37
CALC $NUM20 = $NUM37 / ( 1609.2 ** 2 )
CALC $NUM37 = $NUM20
CALC $NUM39 = ( $NUM37 / $NUM1 ) * 100
DISPLAY $NUM39
ASEL
REM -----
RES MINOR1 = 101
CALC $NUM23 = AREA + $NUM23
CALC $NUM20 = $NUM23 / ( 1609.2 ** 2 )
CALC $NUM23 = $NUM20
CALC $NUM29 = ( $NUM23 / $NUM1 ) * 100
DISPLAY $NUM29
ASEL
REM -----

```

```

RES MINOR1 = 103
CALC $NUM25 = AREA + $NUM25
CALC $NUM20 = $NUM25 / ( 1609.2 ** 2 )
CALC $NUM25 = $NUM20
CALC $NUM27 = ( $NUM25 / $NUM1 ) * 100
ASEL
REM -----
SEL SITES.LAKE
ADD
[translate %sitename%]
[unquote ' ']
[unquote ' ']
[unquote ' ']
[unquote ' ']
[unquote ' ']
[unquote ' ']
[unquote ' ']
[unquote ' ']
[unquote ' ']
[unquote ' ']
[unquote ' ']
ASEL
RES NAME = [quote [translate %sitename%]]
CALC DRAINAGE_AREA = $NUM1
DISPLAY $NUM1
CALC LAKE_AREA = $NUM33
DISPLAY $NUM33
CALC LAKE_PCT = $NUM35
DISPLAY $NUM35
CALC WETLND_AREA = $NUM37
CALC WETLND_PCT = $NUM39
CALC RES_AREA = $NUM23
CALC RES_PCT = $NUM29
CALC SNOW_AREA = $NUM25
CALC SNOW_PCT = $NUM27
Q STOP
    &end    /* for info data block
    &label nocov
    &if [exists temp -cov] &then kill temp
    &end    /* for do
&end    /* for do until
&s cerr [close %fileu%]
&type
&type ALL BASINS COMPLETED ...
&type
&return
/*
&routine get_out
&s cerr [close -all]
&if [exists temp -cov] &then kill temp
&type Program has failed ...
&stop

```



```

/*****
/***** AML PROGRAM CALCPPT.AML *****/
/*****
/* Version:          1.0
/* Language:         Arc Macro Language
/* Arc Version/Platform: 6.1.1 / DG-UNIX
/* Subsystem:        Arc
/*
/*.....
/*
/* Purpose: To populate an INFO file with avgerage annual precipitation
/*          values.
/*
/*.....
/*
/* Arguments: None
/*
/* Before running:
/*   1. INFO file SITES.PRECIP must exist.
/*   2. ASCII file with list of subbasin names for which average annual
/*      precipitation is to be calculated must exist.
/*
/*.....
/*
/* Programs or menus called: none
/*
/*.....
/*
/* History:
/*
/* Author/Site          Date    Version  Event
/* -----
/* HARRISON/USGS-OR    12-28-93   1.0      Original coding
/* BROWNELL/USGS-OR    7-23-94   1.0      Update
/*
/*.....
/*
/* Disclaimer:
/* This program has been used by the U.S. Geological Survey; no warranty, either expressed or
/* implied, is made by the USGS as to the accuracy or functioning of the program.
/* No responsibility is assumed by the USGS in connection with this program's distribution or
/* use.
/*
/*.....
/*
&severity &error &routine get_out
&select [locase [show program]]
&when arc
&do
&type
&type
&type //////////////////////////////////////

```

```

&type
&type Disclaimer:
&type
&type This program has been used by the U.S. Geological Survey;
&type no warranty, either expressed or implied, is made by the USGS as to the
&type accuracy or functioning of the program. No responsibility is assumed
&type by the USGS in connection with this program's distribution or use.
&type
&type //////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////
&type
&type CALCPPT.AML : Version 1.0 : [date -full]
&type
&s tube [extract 1 [show &term]]
&if [null %tube%] &then
  &do
    &term 9999 &mouse
    display 9999 position ul screen
  &end
&end
&otherwise
  &do
    &type This program is run from ARC prompt ...
    &stop
  &end
&end /* select
/*
&if [exists temp -cov] &then kill temp
/*
&if ^ [exists delinlist -file] &then &do
  &type There is no ASCII file (delinlist) containing basin names ...
  &stop
&end
/*
&if ^ [exists sites.precip -info] &then &do
  &type There is no INFO file (sites.precip) to write into ...
  &stop
&end
/*
&if [query 'Purge previous records in SITES.PRECIP'] = .TRUE. &then &do
  &data arc info
ARC
SEL SITES.PRECIP
ASEL
PURGE
Y
Q STOP
  &end
&end
/*
&s fileu [open delinlist oerr -read]
&s rerr = 0
&do &until %rerr% = 102

```

```

&s sitename [read %fileu% rerr]
&if %rerr% ne 102 &then &do
  &type
  &type -----
  &type
  &type Calculating average annual precip for: %sitename%
  &type
  &type -----
  &type
  &if ^ [exists %sitename% -cov] &then &goto nocov
  identity %sitename% precip_net temp poly
&data arc info
ARC
FO $NUM1,4,12,F,3
FO $NUM3,4,12,F,3
FO $NUM5,4,12,F,3
FO $NUM7,4,12,F,3
CA $NUM1 = 0
CA $NUM3 = 0
CA $NUM5 = 0
CA $NUM7 = 0
SEL TEMP.PAT
RES AREA GT 0
REM -- Calculate total area for basin
CA $NUM1 = $NUM1 + AREA
PROGRAM CALCPPT
REM -- For each annual precip polygon
PROGRAM SECTION 2
REM -- Calculate percent of total basin
CA $NUM3 = AREA / $NUM1
REM -- Weight annual precip by percent of total area
CA $NUM5 = $NUM3 * PREC
REM -- Sum weighted precip for basin
CA $NUM7 = $NUM7 + $NUM5
PROGRAM SECTION 3
END
RUN CALCPPT
ERASE CALCPPT
Y
REM -- Add record into file
SEL SITES.PRECIP
ADD
[translate %sitename%]
[unquote]
[unquote]
ASEL
RES NAME CN [quote [translate %sitename%]]
CALC AVG_ANN_PPT = $NUM7
DISPLAY AVG_ANN_PPT
Q STOP
&end /*info data block
&if [exists temp -cov] &then kill temp

```

```
&label nocov
&end /* do
&end /*&do &until
&s cerr [close %fileu%]
&type
&type ALL BASINS COMPLETED ...
&type
&return
/*
&routine get_out
&s cerr [close -all]
&type Program has failed ...
&stop
/*****
/*****
/*****
```

```

/*****
/***** AML PROGRAM ELEV.AML *****/
/*****
/* Version:          1.0
/* Language:         Arc Macro Language
/* Arc Version/Platform: 6.1.1 / DG-UNIX
/* Subsystem:        Arc
/* .....
/*
/* Purpose: To compute mean, minimum and maximum elevation values for drainage areas
/*          located above a selected gaged site.
/* .....
/* Before running: Update list of drainage basin polygon coverage names in line ____ of the
/* program. Be sure that these polgon coverages exist before running the program.
/* .....
/*
/* Programs or menus called: none
/* .....
/* Author/Site      Date      Version  Event
/* -----
/* DARLING/USGS-OR  6-94   1.0      Original coding
/* .....
/* Disclaimer:
/* This program has been used by the U.S. Geological Survey; no warranty, either expressed or
/* implied, is made by the USGS as to the accuracy or functioning of the program.
/* No responsibility is assumed by the USGS in connection with this program's distribution or
/* use.
/* -----
/*
&severity &error &routine get_out
&select [locase [show program]]
&when arc
&do
&type
&type
&type //////////////////////////////////////
&type
&type Disclaimer:
&type
&type This program has been used by the U.S. Geological Survey;
&type no warranty, either expressed or implied, is made by the USGS as to the
&type accuracy or functioning of the program. No responsibility is assumed
&type by the USGS in connection with this program's distribution or use.
&type
&type //////////////////////////////////////
&type
&type ELEV.AML : Version 1.0 : [date -full]
&type
&s tube [extract 1 [show &term]]
&if [null %tube%] &then

```

```

&do
  &term 9999 &mouse
  display 9999 position ul screen
&end
&end
&otherwise
&do
  &type This program is run from ARC prompt ...
  &stop
&end
&end /* select
&severity &error &ignore
GRID
&DO NAME &LIST blaze bullrun cedar clack10 clack20 clack30 clack40 ~
fir lilsan oak10 oak20 salm10 salm20 salm30 salm40 ~
sandy sfbull zig10 zig20
SETCELL 69
SETWINDOW /WTRSHD/SANDY/SITE_COVS/%NAME%
%NAME%GRD = POLYGRID(/WTRSHD/SANDY/SITE_COVS/%NAME%,#,#,#,69)
%NAME%GRD2 = CON(%NAME%GRD > 1,1,0)
ELEV%NAME% = ZONALSTATS(%NAME%GRD2, LATTICE_NET)
KILL %NAME%GRD
KILL %NAME%GRD2
&END
q
&RETURN
&routine get_out
&s cerr [close -all]
&type Program has failed ...
&stop
/*****
/*****
/*****

```